TECHNICAL REPORT NO.

OFFICE OF NAVAL RESEARCH

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DEPARTMENT OF OPERATIONS RESEARCH

AND

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

GENERALIZED PROGRAMMING SOLUTION OF CONTINUOUS-TIME LINEAR-SYSTEM OPTIMAL CONTROL PROBLEMS*

by

G. Stephen Jizmagian

TECHNICAL REPORT NO. 116

January 24, 1969

Supported by the Army, Navy and Air Force under
Contract Nonr-225(53)(NR-042-002)
with The Office of Naval Research

Gerald J. Lieberman, Project Director

*This research was partially supported
by the Office of Naval Research under
Contract Nonr-225(83), NR 373 360

DEPARTMENT OF OPERATIONS RESEARCH
and
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

CONTENTS

ILLUSTRATIONS

TABLE

Chapter I

INTRODUCTION

With the advent of efficient and large-scale mathematical programming techniques, computationally feasible methods are available for optimal control problems. The purpose of this paper is to present an algorithm for solving continuous-time optimal control problems with linear dynamics and various loss criteria. Due to the mathematical programming techniques used in the algorithm, it is well suited for large-scale control problems, i.e., control problems with large numbers of state variables and time-varying control inputs. This work consists of two main results that are combined to develop the algorithm.

In Chapter II, we describe the types of control problems considered, including basic definitions and notations for these problems. The basic results in control theory and certain necessary conditions for optimal control, as described by Pontryagin et al. [1], are also presented.

In Chapter III, the algorithms and basic theorems for linear programming and the simplex method [2], quadratic programming and the complementary pivot theory [3], and the Dantzig-Wolfe generalized program [2] are presented.

The first main result, an algorithm for solving parametric linear and quadratic programming problems, when the objective function is nonlinear in the parameter, is presented in Chapter IV. Also presented is the class of nonlinear functions for which this algorithm is valid. The finiteness of the algorithm, including avoidance of cycling due to degeneracy, is then proven. The characteristics of the optimal solution as a function of the parameter are also described.

The second result, an extension of Dantzig's [9] formulation of optimal control problems as generalized programs, is presented in Chapter V. It is shown that any optimal control problem with the following characteristics may be formulated as a generalized program: (1) the system must initiate from some point in a specified region of the state space; (2) the state at the fixed terminal time can be chosen from another convex region in the state space (fixed initial and final points are included in these definitions); (3) the state of the system is controlled by linear differential equations; (4) the admissible control region is a convex polyhedral set (for each point in time) in the control space; (5) the loss criteria is a linear functional in the state and control and/or a quadratic functional in control and/or the absolute value of the control inputs (minimum fuel), or the minimum time. It is further shown that these continuous-time optimal control problems have an equivalent generalized programming formulation in which the master problem [illegible] linear program o[illegible] or three plus the di[illegible] of the state space. The subproblem to the master program is a parametric programming problem of the control space dimension and is solvable by the methods presented in Chapter IV. This subproblem yields an extreme admissable control that, when used with previously found extreme admissable controls, gives a solution that is closer to a feasible or an optimal one.

The algorithm and its variants are presented in the second part of Chapter V. A flow chart of the algorithm is given, along with a description of each execution. Also included is an initiating phase that terminates in a feasible solution of the control problem. On completion of the initiating phase, the algorithm maintains a feasible control while obtaining new controls; these new controls yield better objective values without disturbing the feasibility. Upper and lower bounds on the optimal objective value are provided at each stage of the algorithm.

In Chapter VI, the characteristics of the optimal controls, without any additional assumptions on the system or on the uniqueness of the solution, are presented. Also included are the relationships between the necessary conditions of Pontryagin and the generalized programming results. Between these optimization conditions, a link exists in the dual variables of the generalized program and the adjoint variables associated with the optimal control problem.

To clarify the algorithm and indicate its computational feasibility, a minimum fuel problem and a minimum time problem are solved in detail in Chapter VII. The convergence properties and solution procedures are illustrated with data obtained from computer runs.

Chapter II

OPTIMAL CONTROL

This section defines an optimal control problem and Pontryagin's necessary conditions for optimality. The emphasis is on those linear systems for which generalized programming equivalents can be formulated.

A. Definition of Dynamic Control Systems

The basic control problem can be described by the differential equations:

$$\dot{x}_i = \frac{dx_i}{dt} = f_i(x_1, \ldots, x_n, u_1, \ldots, u_m, t) \qquad i = 1, 2, \ldots, n , \tag{2.1}$$

where

$$x(t) = \begin{bmatrix} x_1(t) \\ \cdot \\ \cdot \\ \cdot \\ x_n(t) \end{bmatrix} \tag{2.2}$$

is the vector of state variables or phase coordinates which describe the trajectory of the system in Euclidean space through time. The control function is the vector of control inputs

$$u(t) = \begin{bmatrix} u_1(t) \\ \cdot \\ \cdot \\ \cdot \\ u_m(t) \end{bmatrix} \tag{2.3}$$

which influence the state through the differential equations. The system at some initial time, t_o, satisfies the initial conditions,

$$x(t_o) \in S_o \subseteq E^n \ . \qquad (2.4)$$

Thus the system may have an initial point $x(t_o)$ at any one of the points in the set, S_o. At a terminal time, T, the system is required to lie in some region, i.e.,

$$x(T) \in S_T \subseteq E^n \ . \qquad (2.5)$$

The time T may be free or fixed, and the sets, S_o and S_T, may be fixed points.

B. Admissible and Feasible Controls and Reachable Sets

The vector control function, $u(t)$ must be specified at every t and is required to lie in an admissible control region, U_t, where

$$u(t) \in U_t \subseteq E^m \ , \ \forall t \ . \qquad (2.6)$$

Definition 2.1. An admissible control is any vector function, $u(t)$, for which

$$u(t) \in U_t \subseteq E^m \ , \ \forall t \in [0,T] \ ,$$

where $[0,T]$ denotes the time interval $\{t|0 \le t \le T\}$.

The objective of the control problem is to find an admissible control function that transfers the state from some point at t_o to another point at T, while minimizing

$$J = \int_{t_o}^{T} f_o(x_1, \ldots, x_n, u_1, \ldots, u_m, t)\, dt \ . \qquad (2.7)$$

It is convenient at this point to define another variable ,

$$x_o(t) = \int_{t_o}^{t} f_o(x_1, \ldots, x_n, u_1, \ldots, u_m, t)\, dt , \tag{2.8}$$

and to let

$$\bar{x}(t) = \begin{bmatrix} x_o(t) \\ \\ x(t) \end{bmatrix} . \tag{2.9}$$

Thus,

$$\dot{x}_o = f_o(x,u,t) \ , \quad x_o(t_o) = 0 \ , \ \text{and} \quad J = x_o(T) \ . \tag{2.10}$$

Definition 2.2. The reachable set, denoted by R_T, consists of a set of terminal $x(T)$ of admissible solutions to the control problem, without the condition $x(T) \in S_T$.

$R_T = \{x \in E^n | x = x(T)$,

where $x(T)$ is a solution to (2.1) at $t = T$ with

$x(t_o) \in S_o$, $u(t) \in U_t$, $\forall t\}$.

Note that for the fixed final time, T, if

$$S_T \cap R_T = \emptyset ,$$

there is no admissible control to transfer the system from an initial point in S_o to a point in S_T.

Definition 2.3. A control function, $u(t)$, defined for $t \in [t_o, T]$ is a feasible control for the optimal control problem if it is an admissible control and transfers the system from some state $x(t_o) \in S_o$ to a state $x(T) \in S_T$ while $x(t)$ satisfies (2.1). Note that a feasible control exists iff

$$S_T \cap R_T \neq \emptyset .$$

In the optimal control problem, we are searching for a control function, among all feasible controls, that results in a minimal value of J.

Assumption 2.1. We will now restrict our attention to functions f_i, for $i = 0, \ldots, n$, which are autonomous, i.e., they do not depend explicitly on time. We will also assume that the f_i functions for $i = 0, 1, \ldots, n$, are continuous in both x and u and are continuously differentiable with respect to x.

C. The Adjoint System and the Hamiltonian

For any given u or $x(o)$, let $x = x(t)$ be determined by

$$\dot{x}_i = f_i(x, u) \qquad i = 0, 1, \ldots, n .$$

For this choise of u, $x(o)$, and the resulting $x(t)$, we define the adjoint system, $\Psi_0, \Psi_1, \ldots, \Psi_n$, by

$$\dot{\Psi}_i = \frac{d\Psi_i}{dt} = - \sum_{k=0}^{n} \frac{\partial f_k(x,u)}{\partial x_i} \Psi_k$$

$$i = 0, 1, \ldots, n , \tag{2.11}$$

where the partials are evaluated at the above $x(t)$, $u(t)$. The solution to (2.11) is related to the choice of control, $u(t)$.

The <u>Hamiltonian</u> is defined as

$$H(\bar{\Psi},\bar{x},u) = \bar{\Psi}' f(x,u) \quad ,$$

where $\bar{\Psi}'$ is the transpose of

$$\bar{\Psi} = \begin{bmatrix} \Psi_o \\ \Psi_1 \\ \cdot \\ \cdot \\ \cdot \\ \Psi_n \end{bmatrix} ,$$

and

$$f(x,u) = \begin{bmatrix} f_o(x,u) \\ f_1(x,u) \\ \cdot \\ \cdot \\ \cdot \\ f_n(x,u) \end{bmatrix} .$$

Thus (2.1) and (2.11) become

$$\dot{x}_i = - \frac{\partial H}{\partial \Psi_i} \qquad (2.1a)$$

$$\dot{\Psi}_i = - \frac{\partial H}{\partial x_i} \qquad (2.11a)$$

D. <u>Pontryagin's Conditions for Optimality</u>

When the initial and final points, $x(t_o)$ and $x(T)$, are not fixed, the regions S_o and S_T are assumed to be smooth manifolds or convex sets. A necessary condition for optimality in this case is that the solution to (2.1) and (2.11) satisfy a transversality condition.

Let $x(t_o) \in S_o$ and $x(T) \in S_T$ be given points on the boundary of S_o and S_T; and let D_o and D_T be tangent planes of S_o and S_T at these points. Then the solution satisfying (2.1) and (2.11) will also satisfy the transversality condition, if $\Psi(t_o)$ and $\Psi(T)$ are the directions of the supporting hyperplanes, D_o and D_T, of S_o and S_T at $x(t_o)$ and $x(T)$, respectively.

Necessary Conditions for $t_o \leq t \leq T$. Let $u(t)$ be a feasible control with a corresponding trajectory $x(t)$. For $u(t) = u^*(t)$ to yield an optimal solution to the control problem, it is necessary to have a non-zero continuous vector function $\overline{\Psi}(t)$ corresponding to $\overline{x}(t)$ and $u^*(t)$, (2.1) and (2.11), and satisfying the transversality conditions so that

(1) For $t \in [t_o, T]$,

$$H[\overline{x}(t), u^*(t), \overline{\Psi}(t)] = \sup_{u(t) \in U_t} H[\overline{x}(t), u(t), \overline{\Psi}(t)]$$

and

(2) $\Psi_o(T) \leq 0$.

E. The Linear System and Control Constraints

A linear system is defined as a dynamic system in which the $f_i(x_1, \ldots, x_n, u_1, \ldots, u_m)$ are linear in x and u for $i = 1, \ldots, n$. Note that $f_o(x,u)$ need not be linear. This linear system can be described by two matrices, F and G, as

$$\dot{x}(t) = Fx(t) + Gu(t) , \qquad x(t) \in E^n , \quad u(t) \in E^m , \tag{2.12}$$

where F is an $n \times n$ real matrix and G is an $n \times m$ real matrix.

The linear system has a fundamental matrix [4] $e^{F(t-\tau)}$ that has the property of transforming $x(t)$ by:

$$x(t_1) = e^{F(t_1-t_0)} x(t_0) ,$$

when $u(t) = 0$ for $\tau \in [t_0, t_1]$. This fundamental matrix arises from the solution of the differential equations in (2.12) when $u(t) = 0$. The solution for any function $u(t)$ is

$$x(t_1) = e^{F(t_1-t_0)} x(t_0) + \int_{t_0}^{t_1} e^{F(t_1-\tau)} Gu(\tau)\, d\tau . \tag{2.13}$$

When $u(t) \in U_t$ for all t and $x(t_0) \in S_o$, the right-hand side of (2.13) determines a point in the reachable set of U_t, S_o, and time t_1. Hence we can state, for linear systems,

$$R_{t_1} = \left\{ x \in E^n \middle| x = x(t_1) , \right.$$

$$x(t_1) = e^{F(t_1-t_0)} x(t_0) + \int_{t_0}^{t_1} e^{F(t_1-\tau)} Gu(\tau)\, d\tau ,$$

$$\left. u(\tau) \in U_t , \quad \tau \in [t_0, t_1] , \quad x(t_0) \in S_o \right\} .$$

Throughout this paper, we will consider problems where $U_t = U \subseteq E^m$, i.e., the admissible control set is constant over time. We also assume that U is a bounded convex polyhedral set, i.e., it is bounded by hyperplanes in m-dimensional space. Note that any convex polyhedral set can be expressed by

$$U = \{u \in E^m \mid Au \leq b\}$$

for some real fixed matrix A of dimension $q \times m$ and for some real vector b of dimension q.

In the following, we will permit the initial and terminal sets, S_o and $S_T \subseteq E^n$, to be convex sets.

Note that $f_i(x,u) = f_i'x + g_i'u$, where f_i is an n-dimensional vector and g_i is an m-dimensional vector. f_i and g_i are the i^{th} rows of the F and G matrices, respectively.

F. Loss Functionals

In this section, we will describe the different classes of loss functionals. These loss functionals, when combined with linear systems and the above restrictions, can be solved by mathematical programming techniques that are developed and discussed in the next two chapters.

Case 1. Linear Loss Functionals.

We define the linear loss case as one that includes all loss functionals of the form

$$f_o(x,u) = f_o'x + g_o'u ,$$

where f_o and g_o are any real n and m component vectors, respectively. Thus, we can define linear systems with linear loss functionals as completely linear systems.

Case 2. Minimum Fuel Problems.

A certain well-known minimum fuel problem is characterized by loss functionals of the form

$$f(u) = \sum_{i=1}^{m} |u_i| .$$

Case 3. Quadratic Loss in Control.

We consider a function a quadratic only in the control vector, i.e.,

$$f(u) = u'Qu \ ,$$

where Q is a positive semidefinite matrix.

When a linear functional is added to $f(u)$ and modifications of Cases 2 and 3 are permitted, the three cases are:

$$f_o(x,u) = f_o'x + g_o'u + f(u) \ ,$$

$$\text{where} \quad f(u) = \begin{cases} 0 \ , & \text{Case 1} \\ \Sigma|u_i| \ , & \text{Case 2} \\ u'Qu \ , & \text{Case 3} \ . \end{cases}$$

(If $f_o(x,u) = f(u)$, then $f_o = 0$, and $g_o = 0$.) Thus the control problem can be stated as

$$\text{Minimize} \quad x_o(T) \ ,$$

$$\dot{\bar{x}} = \bar{F}\bar{x}(t) + \bar{G}u(t) + f(u)\ U_o \ , \tag{2.14}$$

where

$$\bar{F} = \left[\begin{array}{c|c} 0 & f_o' \\ \hline 0 & \\ \vdots & F \\ 0 & \end{array}\right] ,$$

$$\bar{G} = \left[\begin{array}{c} g_o' \\ \hline G \end{array}\right] ,$$

and

$$U_o = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \in E^n ,$$

$$x(t_o) \in S_o , \quad x_o(t_o) = 0$$

$$x(T) \in S_T , \quad \text{and}$$

$$u(t) \in U = \{u \in E^m | Au \leq b\} .$$

The minimum time problem is also considered where T is not fixed, S_o and S_T are fixed points, and $x_o(T) = T - t_o$.

For linear systems described by matrices F and G and a given polyhedron, U, Pontryagin defines a "general position condition." This condition is satisfied when the vectors Gw, FGw, ..., $F^{n-1}Gw$ are linearly independent in E^n when w has the direction of one of the edges of U. For such systems, at each point of time, t, the function $\Psi(t)'Gu(t)$ achieves its maximum at only one vertex of U, except on a set of measure zero.

Before proceeding further with the development of an algorithm to solve these continuous-time control problems, some of the existing techniques used in solution procedures should be mentioned briefly. Three of these techniques are mentioned here.

<u>Direct Methods [5]</u>. In these methods, admissible and, if possible, feasible controls are chosen to start. The gradient of the cost functional (or, if the starting control is not feasible, a Lagrangian form that takes feasibility into account), with respect to the control function, is determined. Then, by using gradient or steepest descent methods, a new control function is chosen to improve the cost functional (or Lagrangian).

<u>Indirect Methods [6]</u>. Indirect methods primarily seek solutions to the necessary conditions for optimality. Some methods use

arbitrary initial or final conditions for the adjoint variable. In this case, the differential equations, (2.1) and (2.11), are integrated to find solutions for $x(t_o)$; during this procedure, a solution to the necessary conditions is retained, if possible. If $x(t_o)$ is not equal to the original (known) $x^*(t_o)$, the gradient of some cost functional, based on the distance from $x(t_o)$ to $x^*(t_o)$, is used to determine a new guess for the final time adjoint variable values.

Discrete Approximations [7,8]. Mathematical programming techniques, e.g., linear programming or gradient projection methods, are usually applied to a discrete approximation of the continuous-time problem. In these approximations, the system is considered at a prescribed set of instants in the interval $[t_o, T]$. Only at these times are the control inputs allowed to change. The differential equations are then approximated by difference equations for each time considered. Mathematical programming techniques are then used to solve the approximation.

Each of the three techniques mentioned have their disadvantages. The direct methods' disadvantage is that a feasible control must be provided initially. If not, the convergence methods cannot be guaranteed to terminate with a feasible solution. Also, the efficiency of convergence is highly dependent on the initial guess. The indirect methods also have a disadvantage in that they do not provide a feasible solution until the final step. At times, the determination of a feasible solution is the major problem in optimal control. The basic disadvantage of discrete approximations stems from the large number of variables or equations introduced by the approximation process.

The methods developed in this work combine the features of both the direct and indirect methods and use admissible controls to find a feasible solution. This combination continuously reduces the cost while it retains the feasibility and converges on the optimum values of the adjoint variables. Thus the problem, at any iteration in the optimization phase, has a feasible solution available, and the present solution has a measure of closeness to the optimum solution [9].

Chapter III

MATHEMATICAL PROGRAMMING

In this chapter, the algorithms available for solving linear and quadratic programming are reviewed, and the theory of generalized programming is described. The choice of the simplex method for linear programming problems and the complementary pivot theory for quadratic programming problems is dictated by the ease encountered in using the parametric programming methods presented in Chapter IV.

It should be noted that any bounded convex polyhedral set can be represented (possibly after a change of variable) by the set $X = \{x \mid Ax \leq b,\ x \geq 0\}$ for some real matrix, A, and for some real vector, b.

A. Linear Programming

The standard linear programming problem can be stated as

$$\begin{aligned} &\text{minimize} \quad z = c'x \\ &\text{subject to} \quad Ax \leq b \quad \text{and} \\ &\qquad\qquad\quad x \geq 0 \ , \end{aligned} \tag{3.1}$$

where $x \in E^n$, c is a specified n-dimensional vector, b is a specified m-dimensional vector, and A is a given $(m \times n)$ matrix.

Since minimizing $c'x$ is equivalent to maximizing $(-c')x$, only minimization problems are discussed. Hence, problem (3.1) seeks the minimum of a linear (convex and concave) function over a convex polyhedral constraint set; if the latter is nonempty, a solution exists and is known to be at an extreme point in the constraint set. Thus we need only consider basic solutions to problem (3.1), i.e., solutions in which no more than m components of the vector x are positive and whose column coefficients are linearly independent in rows where $Ax \leq b$ is satisfied with equality.

The dual problem to (3.1) can be expressed by

$$\begin{aligned} &\text{minimize} \quad v = by \\ &\text{subject to} \quad A'y \geq c \qquad\qquad (3.2) \\ &\qquad\qquad y \geq 0 \;, \quad y \in E^m \;. \end{aligned}$$

The <u>duality theorem</u> of linear programming can be summarized in two statements:

(1) for any feasible x, y [satisfying the constraints of (3.1) and (3.2)],

$$c'x \geq b'y \;, \quad \text{and}$$

(2) for the optimal x^*, y^* of (3.1) and (3.2),

$$c'x^* = b'y^*$$

$$\left.\begin{aligned} (Ax^* - b)'y^* &= 0 \\ (A'y^* - c)'x^* &= 0 \end{aligned}\right\} \quad \text{complementary slackness conditions} \;.$$

If the x vector is augmented by m components to include slack variables and the matrix A is augmented by I, the constraint inequalities are equivalent to

$$\begin{aligned} Ax &= b \\ x &\geq 0 \;, \end{aligned} \qquad (3.3)$$

where A and x are now the augmented matrix and vector, respectively.

Since we need only investigate the extreme points of the constraint set, we need only allow basic solutions corresponding to choosing m linearly independent columns of A, and the components of the vector x corresponding to the m columns of A. The m columns of the augmented A form a nonsingular matrix B, called the basis matrix. The corresponding components of x are called the basic variables. Hence, a basic feasible solution to (3.1) is one in which the values of the basic variables are nonnegative, and all the other variables, called the non-basic variables, are at value zero. Let x_B represent the vector of the basic variables corresponding to B. Then the basic solution to the linear equations in (3.3) is

$$x_B = B^{-1} b \ ,$$

$$x_i = 0 \ ,$$

where i is nonbasic. This is a basic feasible solution, provided $x_B \geq 0$.

<u>The Simplex Method</u>. The simplex method is reviewed in detail, since a variation of it is employed in Chapter IV for the parametric programming procedures. This method is presented in matrix form. Here, the linear program

$$\begin{aligned} \text{minimize} \quad & z = c'x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{3.4}$$

is observed, and the augmented system of equations

$$\begin{bmatrix} 1 & -c' \\ 0 & A \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \tag{3.5}$$

is used.

Given any basis, B, let the augmented basis be $\overline{B}$, where

$$\overline{B} = \begin{bmatrix} 1 & -c'_B \\ 0 & B \end{bmatrix} ,$$

and rewrite (3.5) as

$$\begin{bmatrix} 1 & -c'_B & -c'_R \\ & & \\ 0 & B & R \end{bmatrix} \begin{bmatrix} z \\ x_B \\ x_R \end{bmatrix} = \begin{bmatrix} 0 \\ \\ b \end{bmatrix} , \tag{3.6}$$

where

$$A = \left[B \;\vdots\; R \right] ,$$

$$x = \begin{bmatrix} x_B \\ \\ x_R \end{bmatrix} , \quad \text{and}$$

$$c = \begin{bmatrix} c_B \\ \\ c_R \end{bmatrix} .$$

c_B is the vector of the components of c corresponding to the basic variables x_B.

Since B is nonsingular, $\overline{B}$ is also nonsingular;

$$\overline{B}^{-1} = \begin{bmatrix} 1 & c'_B B^{-1} \\ & \\ 0 & B^{-1} \end{bmatrix} .$$

Multiplying (3.6) by $\overline{B}^{-1}$ and then rearranging it, we get

$$\begin{bmatrix} z \\ \\ x_B \end{bmatrix} = \begin{bmatrix} 1 & -c'_B B^{-1} \\ & \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ \\ b \end{bmatrix} - \begin{bmatrix} -c'+c'_B B^{-1} R \\ \\ B^{-1} R \end{bmatrix} \begin{bmatrix} x_R \end{bmatrix} . \tag{3.7}$$

By setting the nonbasic variables, x_R, at level zero,

$$\begin{bmatrix} x \\ \\ x_B \end{bmatrix} = \begin{bmatrix} 1 & c'_B B^{-1} \\ \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ \\ b \end{bmatrix} = \begin{bmatrix} c'_B B^{-1} b \\ \\ B^{-1} b \end{bmatrix} = \begin{bmatrix} c'_B x_B \\ \\ B^{-1} b \end{bmatrix},$$

where $x_B = B^{-1}b$.

If $B^{-1}b$ is a nonnegative vector, the basis B is feasible, and thus the current solution is a basic feasible solution.

Look at any variable x_j with a corresponding column A_j and a cost coefficient c_j; this variable's column in the transformed system of (3.7) is

$$\begin{bmatrix} -c_j + c'_B B^{-1} A_j \\ \\ B^{-1} A_j \end{bmatrix} . \qquad (3.8)$$

If x_j is a basic variable, $B^{-1}A_j$ is the r^{th} unit vector, if A_j is the r^{th} column of B. (Note that, in this case, c_j would be the r^{th} component of c_B, and x_j would be the r^{th} component of x_B.) Thus the first component of (3.8) becomes

$$-c_j + c'_B B^{-1} A_j = -c_j + c_j = 0 ;$$

moreover (3.8) is a unit vector.

Proposition 3.1. If all $c_j - c'_B B^{-1} A_j \geq 0$, the current basis B is optimal.

Proof of Proposition 3.1.

Assume $\bar{c}_j = c_j - c'_B B^{-1} A_j \geq 0$, $\forall_j$; then, from (3.7) and (3.8),

$$z = c_B' B^{-1} b + \sum \delta_j x_j \ .$$

Note that $\delta_j = 0$ for j corresponding to a basic variable. Thus any change from the current solution would result in an increase of some x_j (nonbasic), and the value of z would increase or remain unchanged. Hence no improvement in the objective is obtained with any other solution.

Q.E.D.

From Proposition 1, we have an optimality condition for any feasible basis;

$$\delta_j = c_j - c_B' B^{-1} A_j \geq 0 \ , \qquad \forall j \ . \tag{3.9}$$

If, on the other hand, the left-hand side of (3.9) was strictly negative, for some $j = s$, then increasing x_s and adjusting the values of the basic variables until one dropped to value zero (thus replacing a current basic variable) would decrease the objective function, provided x_s entered at a positive level. The simplex method changes the basic set at each iteration with the entering variable, x_s, designated the nonbasic variable with the most negative relative cost factor, δ_j. The exiting variable is the first basic variable to be driven to zero as the entering variable increases above zero (assuming nondegeneracy and bounded solutions). The method terminates with the current basis being optimal, when (3.9) is satisfied for all variables.

When the variable x_s is chosen as the entering variable, the exiting variable can be determined by examining the ratios

$$\frac{\left(B^{-1} b\right)_i}{\left(B^{-1} A_s\right)_i} \ , \qquad \text{for all} \quad i \ , \tag{3.10}$$

where $\left(B^{-1} A_s\right)_i > 0$. From (3.7), the current basic variables are expressed as

$$x_B = B^{-1}b - \left(B^{-1}A_s\right)'x_s \ .$$

Thus the first variable driven to zero in the vector x_B is the one corresponding to the minimum of the ratios defined by (3.10).

The simplex method can be carried out in two ways. The first way (called the revised simplex method) is to substitute A_s which corresponds to the entering variable x_s for A_r which corresponds to the exiting variable x_r in the basis B. With this substitution both the new solution and the relative cost factors can then be calculated. The second way is to pivot in the augmented matrix

$$\begin{bmatrix} -1 & 0 & c_R' - c_B B^{-1} R \\ & & \\ 0 & I & B^{-1}R \end{bmatrix}$$

about the term $\left(B^{-1}R_s\right)_r$, where s corresponds to the entering variable and r corresponds to the exiting variable. The pivoting operations do not change the canonical form of the basic variables which remain basic, but they do force the column

$$\begin{bmatrix} c_s - c_B B^{-1} A_s \\ \\ B^{-1} A_s \end{bmatrix}$$

to the canonical form of

$$\begin{bmatrix} 0 \\ \\ e_r \end{bmatrix} ,$$

where e_r is the unit vector with a one in the r^{th} component; this will alter all of the other columns corresponding to the nonbasic variables. Note that once a feasible basis is determined, the simplex method insures that all succeeding bases are feasible.

To obtain an initial feasible basis, phase I of the simplex method adds artificial variables to (3.4) and solves a new linear program. Let E be an $m \times m$ matrix with only diagonal terms, and let $e_{ii} = +1$, if $b_i \geq 0$, and $e_{ii} = -1$, if $b_i < 0$; then, the new linear program is

$$\min z = \sum_{i=1}^{m} v_i$$

$$Ax + Ev = b$$

$$x \geq 0, \quad v \geq 0 , \tag{3.11}$$

and the solution terminates in a basic feasible solution to (3.4), when the simplex method is applied to (3.11). The optimal value of z in (3.11) is zero iff (3.4) is feasible.

B. Quadratic Programming

The standard quadratic programming [3] problem can be stated as

$$\text{minimize} \quad z = c'x + x'Qx$$

$$\text{subject to} \quad Ax \leq b$$

$$x \geq 0, \tag{3.12}$$

where $x \in E^n$; c is a specified n-dimensional vector; b is a specified m-dimensional vector; A is a specified $(m \times n)$ matrix; and Q is a specified $(n \times n)$ matrix. It is hereby assumed that Q is positive semidefinite.

Since problem (3.12) is a convex programming problem, the Kuhn-Tucker necessary conditions are also sufficient conditions for optimality. Thus a solution, x, to the following necessary conditions is an optimal solution to (3.12).

$$u = c + 2Qx - A'y \geq 0$$

$$v = -b + Ax \geq 0$$

$$x \geq 0$$

$$y \geq 0$$

$$x_i u_i = 0, \quad y_i v_i = 0, \quad \forall i \;. \tag{3.13}$$

If we define

$$w = \begin{bmatrix} u \\ \\ v \end{bmatrix}, \qquad z = \begin{bmatrix} x \\ \\ y \end{bmatrix},$$

$$M = \begin{bmatrix} 2Q & -A' \\ \\ A & 0 \end{bmatrix}, \qquad \text{and} \quad q = \begin{bmatrix} c \\ \\ -b \end{bmatrix},$$

the necessary conditions may be written as

$$\begin{aligned} & w = Mz + q \\ & w, z \geq 0, \qquad w_i z_i = 0, \quad \forall_i = 1, \ldots, p, \end{aligned} \tag{3.14}$$

where M is $p \times p$.

Complementary Pivot Theory. Problem (3.14) is a statement of the fundamental problem of the complementary pivot theory [3]. Although (3.14) is solvable by this theory for various classes of M, the discussion here will be restricted to M being positive semidefinite, as it is in the quadratic programming problem (3.12).

Note that we are looking for a complementary solution to the linear equations in (3.14), i.e., a solution to

$$w = Mz + q \;,$$

with $w_i z_i = 0$, $\forall i$. We will initiate with a solution that is complementary but that may not be nonnegative. We will then retain this complementary property while seeking a nonnegative solution.

The problem in the structured form of

$$\begin{array}{ccccc} & & & & z_1 \quad \cdots \quad z_p \\ \begin{bmatrix} w_1 \\ \vdots \\ w_p \end{bmatrix} & = & \begin{bmatrix} q_1 \\ \vdots \\ q_p \end{bmatrix} & + & \begin{bmatrix} m_{11} & & \sim \\ & \ddots & \\ \sim & & m_{pp} \end{bmatrix} \end{array}, \tag{3.15}$$

is observed with the transformations being made by substituting a variable z_i (or, in later steps, some w_i) in the extreme left column, replacing a variable in the column, and then pivoting on the system of equations by changing the column q and the matrix M. The variables in the left column are called basic, and the variables in the row above the matrix M (or $\overline{M}$ after transformation) are called nonbasic. The problem is initiated by setting $w_i = q_i$ and $z_i = 0$ for all i. If any q_i is negative, pick the w_i corresponding to min q_i, and let it be a distinguished variable. The following can be taken as a general iteration.

> Increase the complement [defined by (3.14)] of the distinguished variable and determine the blocking variable which is either
>
> (a) a basic variable being driven below its lower bound (usually zero) by an increase of the driving variable, or
>
> (b) the distinguished variable which is driven toward zero.
>
> [The first variable to block in either (a) or (b) becomes the blocking variable.]

If the blocking variable is not the distinguished variable, then replace the basic blocking variable with the increasing nonbasic (driving) variable by pivoting about the point $\overline{m}_{rs}$ in the matrix $\overline{M}$, where $\overline{m}_{rs}$ is the term in the current matrix that corresponds to the s^{th} column (the driving variable) and the r^{th} row (the blocking variable). Now increase the complement of the former blocking variable (now nonbasic) until a new blocking variable is found.

If the blocking variable at any stage is the distinguished variable, make it nonbasic at value zero and make the driving variable basic (by pivoting).

At this point, a complementary solution exists. Then look at all $\bar{q}_i$ (determined after pivoting) and choose the most negative to determine the new (basic) distinguished variable. The algorithm terminates when all $\bar{q}_i \geq 0$. The nonbasic variables, placed in the row above the matrix $\bar{M}$, are at level zero, except for the driving variable, at any time.

The pivoting rule is: pivoting on m_{rs},

$$\bar{m}_{rs} = \frac{1}{m_{rs}}$$

$$\bar{m}_{is} = \frac{m_{is}}{m_{rs}}, \qquad \forall i \neq r$$

$$\bar{m}_{rj} = \frac{-m_{rj}}{m_{rs}}, \qquad \forall j \neq s$$

$$\bar{m}_{ij} = m_{ij} - \frac{m_{ir} m_{rj}}{m_{rs}}, \qquad \forall i \neq r$$
$$j \neq s .$$

Let $q_i = m_{i0}$ and apply the pivot rules given above. For basic variables that correspond to negative q_i and not distinguised, we define their common lower bound to be

$$\Gamma < \min q_i ,$$

instead of zero. Thus Γ is the lower bound that blocks the decrease of a basic variable.

It has been showr by Dantzig and Cottle [3] that the algorithm terminates in a solution to the quadratic programming problem when (3.14) has a feasible solution.

C. Generalized Programming

The generalized programming problem can be represented by

Choose a vector, P, in a convex set, $C \subseteq E^n$, such that we

$$\begin{aligned} &\text{maximize} \quad \lambda \\ &\text{subject to} \quad U_0\lambda + P\mu = S \\ &\qquad\qquad \mu = 1, \quad \mu \geq 0 \end{aligned} \tag{3.16}$$

where U_0 and S are specified n-dimensional vectors, and μ is a scalar. [The results here are easily applied to an extended form of (3.16), where the linear equations become

$$U_0\lambda + P_1\mu_1 + P_2\mu_2 + \ldots + P_q\mu_q = S$$

$$\mu_i = 1, \quad \forall i,$$

and each P_i is drawn from a convex set C_i.]

Thus, we are looking for some vector P^* or a convex combination of vector P^{i*}, all in set C, so that the linear equations are feasible, i.e.,

$$U_0\lambda + P^* = S \tag{3.17}$$

or

$$U_0\lambda + \sum_i P^{i*}\mu_i = S$$

$$\sum \mu_i = 1$$

$$\mu_i \geq 0, \tag{3.18}$$

and the resulting value of λ is a maximum over the choice of all the elements in set C, which satisfy the linear equations. Note that, if any set of P^i is in C, any convex combination of that set is also in C. Hence (3.17) and (3.18) are equivalent when

$$P^* = \sum_i P^{i*}\mu_i$$

$$\sum \mu_i = 1, \quad \mu_i \geq 0 .$$

The solution procedure assumes we have on hand, initially, n particular choices of $P^i \in C$ so that the following linear program (called a restricted master)

$$\max_{\mu} \quad \lambda$$

$$\text{subject to} \quad U_0\lambda + P^1\mu_1 + \ldots + P^n\mu_n = S$$

$$\mu_1 + \ldots + \mu_n = 1$$

$$\mu_i \geq 0 \qquad (3.19)$$

has a unique, feasible, nondegerate solution with the basis being defined as

$$B^0 = \begin{bmatrix} U_0 & P^1 & \ldots & P^n \\ & & & \\ 0 & 1 & \ldots & 1 \end{bmatrix}$$

and being nonsingular (by definition). Since for each $P^i \in C$, $P^0 = \sum P^i\mu_i^0$, where μ_i^0 is a solution to (3.19), is in C and is a feasible solution to (3.16), but not necessarily the optimal solution.

To test P^0 [and hence, any solution to (3.16), generated from a basis] for optimality, a row vector $\bar{\pi} = \bar{\pi}^0$ is determined to satisfy

$$\bar{\pi}^0 B^0 = (1, 0, \ldots, 0) \; . \qquad (3.20)$$

From $\bar{\pi}$, we find a vector P^{n+1}, which is not necessarily unique, and a value δ so that

$$\delta = \bar{\pi}^0 \bar{P}^{n+1} = \min_{P \in C} \bar{\pi}^0 \bar{P} , \tag{3.21}$$

where $\bar{P} = \begin{bmatrix} P \\ 1 \end{bmatrix}$. If $\delta = 0$, the current solution is an optimal one. If $\delta < 0$, (3.19) is augmented by P^{n+1} and the new linear program is then solved. The general iteration starts with a solution to the restricted <u>master program</u>

$$\begin{aligned} &\text{maximize} \quad \lambda \\ &\text{subject to} \quad U_0 \lambda + \sum_{1}^{n+k} P^i \mu_i = S \\ &\qquad\qquad\qquad \sum \mu_i = 1 \\ &\qquad\qquad\qquad \mu_i \geq 0 . \end{aligned} \tag{3.22}$$

Let B^k be the optimal basis to the linear program (3.22), and let $\bar{\pi}^k$, the dual (optimal) variable to (3.22), be defined analogous to (3.20). Then, δ^k and P^{n+k+1} are found from the <u>subproblem</u>,

$$\text{find} \quad \delta^{k+1} = \bar{\pi}^k \bar{P}^{n+k+1} = \min_{P \in C} \bar{\pi}^k \bar{P} . \tag{3.23}$$

If $\delta^{k+1} = 0$, the solution to the k^{th} iteration of the master problem is optimal. If $\delta^{k+1} < 0$, then P^{n+k+1} can be adjoined to (3.22), and the solution to (3.16) is improved. The value $-\delta^{k+1}$ is the maximum amount by which the value of the current basis λ^k can be improved. Thus, $\lambda - \delta^{k+1}$ constitutes an upper bound to the optimal solution of (3.16). It is known that these upper bound evaluations can vary considerably from one interation to the next. Accordingly, the least of these evaluations is saved from all iterations, including the current one.

It has been shown that, if C is bounded and the initial solution to (3.19) is nondegenerate ($\mu_i > 0$), $\bar{\pi}^k \to \bar{\pi}^*$ and $P^{k*} \to P^*$ [where

$$P^{k*} = \sum_{i=1}^{n+k} P^i \mu_i ,$$

and μ_i is a solution to (3.22)], on a subsequence k, and that $P = P^*$ is optimal for (3.16). $\bar{\pi}^*$ satisfies the properties

$$\bar{\pi}^* \neq 0$$

$$\bar{\pi}^* \bar{P} \geq \bar{\pi}^* \bar{P}^* = 0, \qquad \text{for all } P \in C . \tag{3.24}$$

Moreover, if C is a polyhedral set, then the subproblem (3.23) is a linear program, and the iterative process terminates in a finite number of steps. It should be noted that, in any case, the objective function improves with each iteration, and a feasible solution always exists to the master problem. Also, the initial solution (or columns) for (3.19) can be obtained by a procedure similar to a phase I simplex method.

Remembering that the usual form of a generalized program includes the sum of the vectors $P_i \in C_i$, where the C_i are convex sets, the vector S need not be fixed, but it must be drawn from a convex set, $\mathcal{S}$. Thus the generalized program becomes

$$\max_{P,S}$$

$$U_0 \quad P_i - S = 0$$

$$= 1$$

$$= 1 , \tag{3.25}$$

where $P \in C$ and $S \in \mathcal{S}$. In this case, the subproblem is extended to

find $\ ^k$ as in (3.23) and

$$\Delta^k = \min_{S \in \mathcal{S}} \bar{\pi}^k \bar{S},$$

where $\bar{S} = \begin{bmatrix} S \\ i \end{bmatrix}$. If δ^k or $\Delta^k < 0$, then the corresponding vector or vectors is entered into the master problem. If both δ^k and $\Delta^k = 0$, the current solution is optimal.

The generalized programming problem,

Primal: $\max_{P} \lambda$

$$U_0 \lambda + P\mu = S$$

$$\mu = 1$$

$$P \in C \qquad (3.16)$$

has as its

Dual: find a vector $\bar{\pi}$ so that

$$\bar{\pi} \begin{bmatrix} P \\ 1 \end{bmatrix} \geq 0, \qquad \forall P \in C$$

$$\bar{\pi} \begin{bmatrix} P \\ 1 \end{bmatrix} = 0, \qquad \text{some } P \in C$$

$$\bar{\pi} \begin{bmatrix} U_0 \\ 0 \end{bmatrix} = 1. \qquad (3.26)$$

This dual is the equivalent of finding a particular hyperplane to support the convex set C. If a solution to the dual is known, then a solution to the primal may be found using the dual solution, $\bar{\pi}^*$, to find the vectors, $P^* \in C$, that satisfy

$$\bar{\pi}^* \begin{bmatrix} P^* \\ 1 \end{bmatrix} = 0.$$

If P^* is unique and the primal has a solution, P^* must be the solution. If P^* is not unique and the primal has a solution, then some convex combination of all the P^* must form the primal solution.

Chapter IV

PARAMETRIC PROGRAMMING

In this chapter, algorithms are presented for solving parametric linear and quadratic programming problems, where the dependence on the parameter is nonlinear and occurs only in the linear part of the objective function. These parametric programming problems arise in the subproblem of the generalized programming formulation of the optimal control problems.

A. Parametric Linear Programming

We consider the following problem linear in x

$$\begin{aligned} &\text{find } x^*(t) \text{ to} \\ &\text{minimize } \gamma(t)'x \\ &\text{subject to } Ax = b \\ &\qquad x \geq 0 \\ &\qquad t \in [T_1, T_2] \,, \end{aligned} \tag{4.1}$$

and the following problem quadratic in x

$$\begin{aligned} &\text{find } x^*(t) \text{ to} \\ &\text{minimize } \gamma(t)'x + xQx \\ &\text{subject to } Ax \geq b \\ &\qquad x \geq 0 \\ &\qquad t \in [T_1, T_2] \,. \end{aligned} \tag{4.2}$$

In both of the above cases, A is a given $m \times n$ real matrix, b is a given n-dimensional vector, Q is an $n \times n$ positive semidefinite matrix, x is a vector in E^n, and

$$\gamma(t) = [\gamma_1(t), \ldots, \gamma_k(t), \ldots, \gamma_N(t)] \tag{4.3}$$

is a given vector valued function, each component of which is a solution to some homogeneous, linear differential equation with constant real coefficients that may depend on k. Such $/_k(t)$ are of the form

$$/_k(t) = \sum_{i=1}^{N} p_{ki}(t)\, e^{s_{ki} t}, \qquad (4.4)$$

where $p_{ki}(t)$ is a polynomial with real coefficients of degree m_{ki} so that

$$\sum_{i=1}^{N} m_{ki} = N$$

and s_{ki} are constants so that, if s_{ki} is complex for i odd, s_{ki+1} is its conjugate and $p_{ki}(t) = p_{ki+1}(t)$. It follows then that these $/_k(t)$ are real-valued functions of t.

The lemmas and theorems that follow are required to show convergence of the proposed algorithm.

<u>Lemma 4.1.</u> If $/_*(t)$ is a solution to a homogeneous linear differential equation with constant real coefficients of order N and if for some $t = t_0$

$$/_*(t)\Big|_{t_0} = \frac{d}{dt}\, /_*(t)\Big|_{t_0} = \dots = \frac{d^{N-1}[/_*(t)]}{dt^{N-1}}\Big|_{t_0} = 0$$

then $/_*(t) \equiv 0$ for all t.

<u>Proof of Lemma 4.1.</u>

$/_*(t)$ solves an equation of the form

$$\frac{d^N}{dt^N}\gamma_*(t) = a_0\gamma_*(t) + a_1\frac{d\gamma_*(t)}{dt} + \ldots + a_{N-1}\frac{d^{N-1}\gamma_*(t)}{dt^N} . \tag{4.5}$$

At $t = t_0$,

$$\left.\frac{d^N}{dt^N}\gamma_*(t)\right|_{t=t_0} = 0 . \tag{4.6}$$

By taking the derivative of both sides of (4.5) and substituting (4.6),

$$\left.\frac{d^{N+1}}{dt^{N+1}}\gamma_*(t)\right|_{t=t_0} = 0 .$$

If this procedure is continued, all derivatives of $\gamma_*(t)$ at $t = t_0$ become zero. Therefore, with $\gamma_*(t) = 0$ and all of its derivatives at zero for $t = t_0$ and with $\gamma_*(t)$ being able to expand (at $t = t_0$) to a Taylor series, $\gamma_*(t)$ must be constant and have value zero for all t.

Q.E.D.

<u>Definition 4.1.</u> A vector y is said to be lexicographically greater than zero, if at least one component is non-zero and the first such component is positive; this vector can be denoted as

$$y \succ 0 .$$

A vector y is lexicographically greater than a vector z,

$$y \succ z$$

if $y - z \succ 0$. A vector is said to be lexicographically greater or equal to zero, if it is lexicographically greater than zero or equal to zero.

A similar definition is true for one vector to be lexicographically greater or equal to another vector.

Definition 4.2. Let $\gamma_*(t)$ be a real scalar function of t. Then the N-component vector $D_{\gamma_*}(t)$ can be defined by its components

$$D_{\gamma_*}(t)_i = \frac{d^{i-1}\gamma_*(t)}{dt^{i-1}}$$

Thus the vector function $D_{\gamma_*}(t)$ is defined by the function $\gamma_*(t)$ and its first $N-1$ derivatives.

Lemma 4.2 [10]. If $f(x)$ has a derivative at c and $f'(c) > 0$, then a positive number δ exists so that for $c < x < c + \delta$, $f(c) < f(x)$.

Theorem 4.1. Let $\gamma_*(t)$ be a member of the class of solutions to homogeneous, constant real coefficient, N^{th} order, linear differential equations, and let $D_{\gamma_*}(t)$ exist as it is defined in Definition 4.2. Then, if $D_{\gamma_*}(t_0) = 0$ or if $D_{\gamma_*}(t_0) \succ 0$, a $\delta > 0$ exists so that $\gamma_*(t) \geq 0$ when $t \in [t_0, t_0 + \delta)$.

Proof of Theorem 4.1.

If $D_{\gamma_*}(t_0) = 0$, then, according to Lemma 4.1, $\gamma_*(t) = 0$ for all t and $\delta = \infty$.

If $D_{\gamma_*}(t_0) \succ 0$, either $\gamma_*(t_0) > 0$ or its lowest order derivative--one that is non-zero at $t = t_0$--is greater than zero. If $\gamma_*(t_0) > 0$, then, by continuity a $\delta > 0$ exists for $\gamma_*(t_0) \geq 0$ when $t \in [t_0, t_0 + \delta)$. If $\gamma_*(t_0) = 0$, let the lowest order, non-zero derivative at $t = t_0$ be the j^{th} derivative, and let

$$f(t) = \frac{d^{j-1}}{dt^{j-1}} \gamma_*(t) .$$

Thus $f'(t_0) > 0$ and, by using Lemma 4.2, a δ does exist for $t_0 < t < t_0 + \delta$, so that

$$\frac{d^{j-1}}{dt^{j-1}} \gamma_*(t_0) = f(t_0) < f(t) = \frac{d^{j-1}}{dt^{j-1}} \gamma_*(t) .$$

From Taylor's theorem, it is known that a number τ, between t_0 and t, exists for any given $t, t_0 < t < t_0 + \delta$ and

$$\gamma_*(t) = \gamma_*(t_0) + \frac{d}{dt} \gamma_*(t_0)(t-t_0) + \ldots + \frac{d^{j-1}}{(j-1)!\,dt^{j-1}} \gamma_*(\tau)(t-t_0)^{j-1} .$$

Since

$$\gamma_*(t_0) = \frac{d}{dt} \gamma_*(t_0) = \ldots = \frac{d^{j-2}}{dt^{j-2}} \gamma_*(t_0) = 0$$

and

$$\frac{d^{j-1}}{dt^{j-1}} \gamma_*(\tau) > \frac{d^{j-1}}{dt^{j-1}} \gamma_*(t_0) = 0 ,$$

$\gamma_*(t) \geq 0$. Hence, it follows that $\gamma_*(t) \geq 0$ for all $t \in [t_0, t_0 + \delta)$.

Q.E.D.

The first algorithm presented here is based on the simplex method and solves problem (4.1).

Find $x^*(t)$ to

minimize $\gamma(t)'x$

subject to $Ax = b, x \geq 0, \quad t \in [T_1, T_2]$.

Here, $\gamma(t)$ is a vector whose i^{th} component, $\gamma_i(t)$, is a real scalar function. At time t_0, let B_0 be an optimal basis for the linear program

$$\begin{aligned} \text{minimize} \quad & \gamma(t_0)'x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \; . \end{aligned} \tag{4.7}$$

Let the solution be $x = x^*(t_0)$. Since $x^*(t_0)$ is a feasible solution to $Ax = b, x \geq 0$, it remains a feasible one to problem (4.1) for all t, but it is not necessarily an optimal solution. Thus how the optimality test for B_0 varies as t takes on the values $t = t_0 + \epsilon$, where $\epsilon > 0$, must be investigated. Let

$$\bar{\gamma}_k(t, B_0) = \gamma_k(t) - \gamma_{B_0}(t)\, B_0^{-1} A_k \, , \tag{4.8}$$

where A_k is the k^{th} column of the matrix A,

$$\epsilon_k = \inf\{\epsilon \mid \bar{\gamma}_k(t_0 + \epsilon, B_0) < 0\} \, , \qquad \text{and} \tag{4.9}$$

$$\epsilon_0 = \min_k \epsilon_k \; . \tag{4.10}$$

It is possible that the ϵ_0, presented above, is zero. $\bar{\gamma}_k(t, B_0)$ is the relative cost factor for any t of column k when B_0 is chosen as the basis. Then the ordering of columns A can be taken so that $A_1, \ldots, A_m$ correspond to the m columns of B_0.

For any basic variable x_i, associated with the optimal basis B_0,

$$\bar{\gamma}_i(t, B_0) = \gamma_i(t) - \gamma_{B_0}(t)\, B_0^{-1} A_i = \gamma_i(t) - \gamma_i(t) = 0, \quad \forall t \; . \tag{4.11a}$$

And, for $t = t_0$,

$$\bar{\gamma}_k(t_0, B_0) \geq 0, \quad \forall k \, , \tag{4.11b}$$

with equality being held when k corresponds to a basic variable. Therefore, the solution to (4.1) remains optimal, i.e., it satisfies $\bar{\gamma}_k(t, B_0) \geq 0$ when $t \in [t_0, t_0 + \epsilon_0)$ for some $\epsilon_0 > 0$, given by (4.10).

Let $\bar{\gamma}^j(t, B)$ be the j^{th} derivative (with respect to t) of the relative cost vector $\bar{\gamma}(t, B)$ when B is the basis under consideration. Let $\bar{\gamma}_k^j(t, B)$ be the component of the above vector corresponding to column k of A. Let A^1 be the new linear programming matrix obtained after deleting all columns (variables) for which the relative cost factors $\bar{\gamma}_k(t_0, B_0) = \bar{\gamma}_k^0(t_0, B_0)$ are strictly positive for $t = t_0$. A general iteration is given with $t_0 = T_1$; a flow chart of the algorithm follows the iteration.

Step I: Solve the linear program

$$
\begin{aligned}
\text{minimize} \quad & c(t_0)'x \\
\text{subject to} \quad & Ax = b \\
& x \geq 0
\end{aligned}
\tag{4.12}
$$

to obtain the optimal basis B_0. If the solution is unique at t_0^+ (i.e., all relative cost factors for nonbasic variables are strictly positive), proceed to Step III.

If the solution is non-unique, fix $t = t_0$ and proceed to Step II, starting with $j = 1$ and $A^0 = A$.

Step II: Let the matrix A^j be composed of the matrix B_{j-1} and all k columns of A^{j-1} having the relative cost factors $\bar{\gamma}^{j-1}(t_0, B_{j-1}) = 0$, where $\bar{\gamma}^{j-1}(t_0, B_{j-1})$ refers to the $(j-1)^{st}$ derivative. To simplify notation in (4.13) below, let the new x and γ vectors corresponding to A^j also be denoted by x and γ, although they are now shortened x and γ vectors. Then, solve the linear program

$$\text{minimize} \quad \bar{\gamma}^j(t_0, B_{j-1})'x$$
$$\text{subject to} \quad A^j x = b$$
$$x \geq 0 \,. \tag{4.13}$$

Let B_j denote the optimal basis. [Computationally, it is convenient to start with the previously optimal, basic feasible solution corresponding to $j-1$ and then to apply the simplex method to obtain an optimal solution to (4.13).] If the solution to (4.13) is unique or if $j = N - 1$, use the optimal basis B_j and proceed to Step III.

If the solution is non-unique and $j < N - 1$, increase j by 1 and repeat Step II.

Step III: Using the optimal basis from Step I or II in place of B_0 in (4.8) and (4.9) for all columns k (optimal basic columns can be ignored since their $\epsilon_k = +\infty$) find ϵ_k. Then calculate ϵ_0. The solution, $x^*(t) = B^{-1}b$, is then optimal for all $t \in [t_0, t_0 + \epsilon_0]$. Moreover, it will be shown that $\epsilon_0 > 0$.

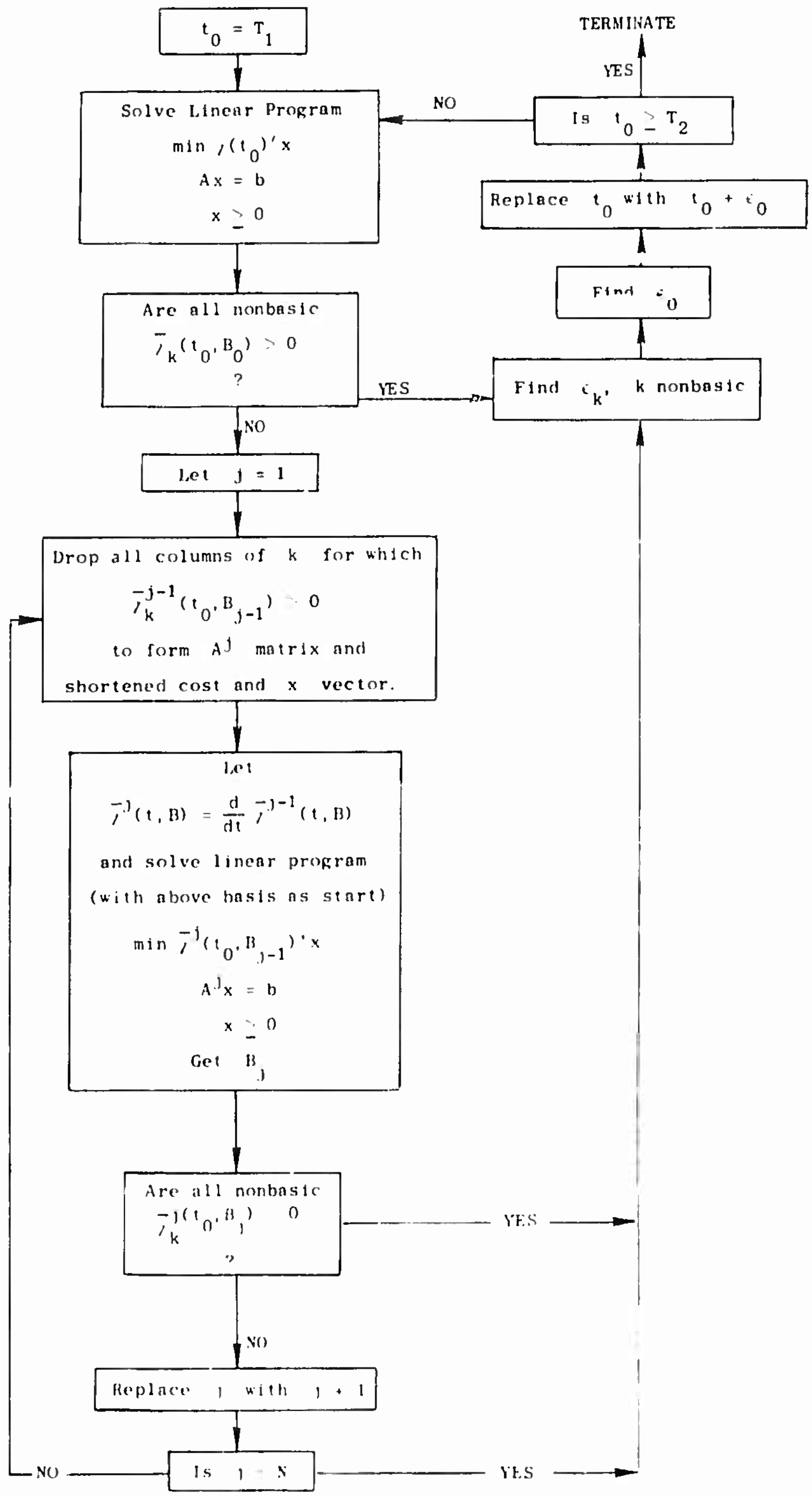
If $t_0 + \epsilon_0 \geq T_2$, the parametric programming problem is solved. However if the solution is not reached, repeat the general iteration with $t_1 = t_0 + \epsilon_0$ replacing t_0.

That this algorithm does terminate in a finite number of steps to a solution of (4.1) for all $t \in [T_1, T_2]$ remains to be shown. The remainder of this section is devoted to showing a finite number of steps to the solution.

Lemma 4.3. If the relative cost factors for some basic B_0 are zero for a subset of columns S and positive for the remaining columns T, then the same is true for any basis B_1 whose columns are in S.

Proof of Lemma 4.3.

The vector of coefficients of the objective equation of the original matrix can be replaced by the relative cost vector for B_0. The price vector π^0 of simplex multipliers relative to B_0 satisfies $\pi^0 B_0 = 0$



FLOW CHART FOR PARAMETRIC LINEAR PROGRAMMING ALGORITHM and that of $\pi^1 B_1 = 0$, since the objective coefficients for both columns of B_0 and B_1 are now zero by hypothesis. Hence $\pi^0 = \pi^1 = 0$, and it follows that their relative cost factors are identical.

Q.E.D.

<u>Theorem 4.2.</u> At the end of the general iteration, the following vector relations are satisfied:

$$D[\bar{\gamma}_k(t_0, B)] = \begin{bmatrix} \bar{\gamma}_k(t_0, B) \\ \frac{d}{dt}\bar{\gamma}_k(t_0, B) \\ \vdots \\ \frac{d^{N-1}}{dt^{N-1}}\bar{\gamma}_k(t_0, B) \end{bmatrix} \succ 0$$

for $k = 1, \ldots, n$

or

$$D[\bar{\gamma}_k(t_0, B)] = 0 \qquad \text{for all } k,$$

where B is the final basis on terminating the iteration of Step II at t_0.

<u>Proof of Theorem 4.2.</u>

For k corresponding to the basic variables of (4.11), $D[\gamma_k(t_0, B)] = 0$. For the basis B, let the basic variables be x_B, and let the nonbasic variables be x_R. The problem is separated as

$$
\begin{bmatrix} B & \vdots & R \end{bmatrix} \begin{bmatrix} x_B \\ x_R \end{bmatrix} = b
$$

$$
z - \begin{bmatrix} \gamma_B', \gamma_R' \end{bmatrix} \begin{bmatrix} x_B \\ x_R \end{bmatrix} = 0 \; . \tag{4.14}
$$

By pivoting on the $m + 1$ rows of (4.14), the first m columns are now unit vectors, and the system becomes

$$
\begin{bmatrix} I & \vdots & \bar{R} \end{bmatrix} \begin{bmatrix} x_B \\ x_R \end{bmatrix} = \bar{b}
$$

$$
z - \begin{bmatrix} 0 & \vdots & \bar{\gamma}_R' \end{bmatrix} \begin{bmatrix} x_B \\ x_R \end{bmatrix} = \bar{z}_0 \; .
$$

where $\bar{\gamma}_R$ is the relative cost factor vector of the nonbasic variables. $\bar{R}$ is separated into two matrices, $\bar{R}_1$ and $\bar{R}_2$, so that the relative cost factors corresponding to the columns of $\bar{R}_1$ are zero, and those corresponding to the columns of $\bar{R}_2$ are negative. Then, the problem

$$
\begin{array}{ll}
\text{rows } 1, \ldots, m & \begin{bmatrix} I & \vdots & \bar{R}_1 & \vdots & \bar{R}_2 \end{bmatrix} \begin{bmatrix} x_B \\ x_{\bar{R}_1} \\ x_{\bar{R}_2} \end{bmatrix} = \bar{b} \\[2em]
\text{row } m+1 & z - \begin{bmatrix} 0 & \vdots & 0 & \vdots & \bar{\gamma}_{\bar{R}_2}(t_0, B) \neq 0 \end{bmatrix} \begin{bmatrix} x_B \\ x_{\bar{R}_1} \\ x_{\bar{R}_2} \end{bmatrix} = \bar{z}_0
\end{array}
$$

$$\begin{array}{l} \text{row} \\ m+2 \end{array} \qquad \min z_1 - \left[\, 0 \;\vdots\; \bar{R}_1^{-1'} (t,B) \;\vdots\; \sim \right] \begin{bmatrix} x_B \\ x_{\bar{R}_1} \\ x_{\bar{R}_2} \end{bmatrix} = 0 \qquad (4.15)$$

is observed, where only the variables corresponding to the columns of $\bar{R}_1$ and I are allowed to enter the basis.

According to Lemma 4.3, the pivoting procedures of the simplex algorithm retain the zero elements of row $m + 1$ at level 0 and the positive elements of row $m + 1$ are at positive values for every stage; these procedures terminate with all relative cost factors of row $m + 2$, corresponding to I and $\bar{R}_1$, at nonnegative values. At termination, because of the simplex method stopping rule, a new set of basic variables is found having the property of the component of the basic variables in rows $m + 1$ and $m + 2$ being at zero (after pivoting); the components of the nonbasic variables are either

(1) zero in row $m + 1$ and nonnegative in row $m + 2$, for variables corresponding to columns of I or $\bar{R}_1$ in (4.15) or

(2) strictly positive in row $m + 1$, for variables corresponding to columns of $\bar{R}_2$.

If the variables are as in (1) above, the ones having zero components in row $m + 2$ are chosen with their columns for consideration in the next stage of the algorithm. Once a nonbasic variable has a positive relative cost factor at any stage j, it can be assumed that its relative cost factors were at zero in previous stages also; hence it can no longer enter the basis. Since the relative cost factors in the first $k - 1$ stages can never change sign by pivoting in the k^{th} and later stages, its derivative vector must be lexicographically greater than 0. Thus at completion of the N stages, all derivative vectors must be lexicographically greater than or equal to zero.

Q.E.D.

<u>Theorem 4.3</u>. The basis B, obtained at the end of Step II in the algorithm for any $t = t_0$, remains optimal for the interval $[t_0, t_0 + \epsilon_0]$, where ϵ_0 is strictly positive.

Proof of Theorem 4.3.

Since each $\bar{\gamma}_k(t,B)$ is a member of the class of solutions to homogeneous, constant coefficient, linear differential equations, ϵ_k is strictly positive by Theorem 4.1. By the definitions given for ϵ_k and t_0, the basis B satisfies the optimality criteria for

$$t \in [t_0, t_0 + \epsilon_0] .$$

Q.E.D.

Theorem 4.4. The number of basis changes in any finite interval $[T_1, T_2]$ is a finite number, and the parametric programming problem is solvable in a finite number of steps.

Proof of Theorem 4.4.

At any switching point t_0, there exists a basis B_p and an $\epsilon_p > 0$, so that B_p is optimal for $t \in [t_0, t_0 + \epsilon_p)$. There also exists a basis B^* and an $\epsilon^* > 0$, making B^* optimal for $t \in [t_0 - \epsilon^*, t_0)$. It follows then, if there is a cluster point at t_0, there would be an infinite increasing sequence of switching points $t_i \in [t_0 - \epsilon^*, t_0]$ which could be bypassed by a single switch at any such t_i to basis B^*. This establishes the existence of a finite number of basis changes in any finite interval $t \in [T_1, T_2]$.

What remains to be shown is that the algorithm, as presented, solves the parametric programming problem in a finite number of steps. As discussed above, let us assume there is a switching point $t_i \in [t_0 - \epsilon^*, t_0]$ and a switch from basis B_{i-1} to B_i. To simplify the discussion, let us also assume that B_i differs from B_{i-1} by the introduction of one incoming column k, and that the value of the incoming variable $x_k = x_k^0 > 0$, i.e., the basic solution is nondegenerate.

The optimal value of the objective z in the neighborhood of t_i, takes the form of

$$z = z_0(t, B_{i-1}) + \bar{\gamma}_k(t, B_{i-1})' x_k , \qquad x_k = \begin{cases} 0, & \text{if } t \leq t_i \\ x_k^0 > 0, & \text{if } t > t_i \end{cases}$$

Both $z_0(t,B_{i-1})$ and $\bar{\lambda}_k(t,B_{i-1})$ can be shown to be linear combinations of solutions to homogeneous, linear differential equations with constant real coefficients and hence they themselves are also solutions. In addition for t_i^+, $\bar{\lambda}_k(t_i^+,B_{i-1}) < 0$, otherwise there would have been no switch from B_{i-1} to B_i. This means that the vector of the 0, 1^{st}, 2^{nd}, ..., $N\text{-}1^{st}$ order derivatives of $\bar{\lambda}_k(t,B_{i-1})$, evaluated at t_i, is lexicographically negative. It follows then that the function z is <u>discontinuous</u> in t at t_i in at least one of its j^{th} order derivatives. On the other hand, the optimal value of z, in the interval $t \epsilon [t_0 - \epsilon^*, t_0]$, can also be expressed as

$$z = z_0(t,B^*)$$

and is continuous in all derivatives at t_i, which is a contradiction.

The above argument is now extended to the case where B_{i-1} is assumed to differ from B_i by several incoming columns. The term $x_k\bar{\lambda}_k(t,B_{i-1})$ is then replaced by a sum of terms, each of which is lexicographically negative at $t = t_i$, hence their sum is lexicographically negative and the discontinuity at $t = t_i$ follows. (Note that the degeneracy of basic solutions in the simplex algorithm is assumed to be handled by the standard right-hand side lexicographic rules of the simplex method.)

Since each basis change is accomplished by a finite (at most N) number of linear programs, the parametric programming problem is solvable in a finite number of steps.

Q.E.D.

<u>Corollary 4.1.</u> The solution $x^*(t)$ to (4.1) is a piecewise constant vector function of t with a finite number of discontinuities.

<u>Proof of Corollary 4.1.</u>

Since $x^*(t)$ is constant for $t \epsilon [t_0, t_0 + \epsilon_0]$, this follows immediately from Theorem 4.4.

Q.E.D.

Proposition 4.1. When all the s_j for each $\gamma_i(t)$ are real, the upper bound on the number of switchings in an infinite time interval for an $(m \times m)$ matrix A is

$$\binom{n}{m}(n - m)\ N\ .$$

Proof of Proposition 4.1.

For any $\gamma_*(t)$ that is a member of the class of solutions to an N^{th} order homogeneous, constant coefficient, linear differential equation and for real and distinct s_i, it is known [1] that the function $\gamma_*(t)$ has, at most, N roots of $\gamma_*(t) = 0$.

At the most, there are $\binom{n}{m}$ possible bases for A; for each of these bases, there are (n-m) nonbasic variables. For each nonbasic variable x_i, the relative cost factor $\bar{\gamma}_j(t,B)$ has at most N points at which it crosses the value zero and thereby creates a possible basis switch.

Q.E.D.

B. Parametric Quadratic Programming

The conditions for which the parametric programming problem has a solution are found in this section. Also, an algorithm based on the complementary pivot theory procedure for quadratic programming is constructed.

The quadratic programming problem

$$\begin{aligned} &\text{find } x^*(t) \text{ to} \\ &\text{minimize } \gamma(t)'x + xQx \\ &\text{subject to } Ax \geq b \\ &\qquad x \geq 0\,, \qquad t \in [T_1, T_2]\,, \end{aligned} \tag{4.16}$$

can be formulated in the complementary pivot theory as,

$$\text{find} \quad w^*(t), \quad z^*(t), \quad \text{so that}$$

$$w = mz + q(t)$$

$$w_i z_i = 0, \quad \forall i$$

$$w_i \geq 0, \quad z_i \geq 0, \quad \forall i \, , \tag{4.17}$$

where

$$M = \begin{bmatrix} 2Q & -A' \\ A & 0 \end{bmatrix}, \qquad q(t) = \begin{bmatrix} \gamma(t) \\ -b \end{bmatrix},$$

$$z = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \text{and} \qquad w = \begin{bmatrix} u \\ v \end{bmatrix},$$

where y is the vector of dual variables to the quadratic programming problem, and u, v are slack vectors of the necessary conditions for quadratic programs, as was stated in Chapter III. The necessary conditions in (4.17) for Q positive semidefinite are sufficient at any $t = t_0$. From the results of Dantzig and Cottle [3], the complementary pivot theory algorithm terminates in a solution to (4.17) when M is positive semidefinite, providing the solution set for

$$w = Mz + q, \qquad w \geq 0, \qquad z \geq 0,$$

is nonempty.

To show that the above solution set is nonempty, a solution must be shown to exist for every t in the parametric programming problem. If such a solution does exist, it must satisfy the conditions of (4.17). Therefore, let us assume the absence of unbounded solutions, i.e., that the set

$$X = \{x \mid x \geq 0,\ Ax \geq b\}$$

is bounded and nonempty.

<u>Proposition 4.2.</u> The parametric quadratic programming problem has a solution for every point $t \in [T_1, T_2]$, when X is nonempty and bounded, $\gamma(t)$ is a vector function with each of its components bounded in the interval $[T_1, T_2]$, and Q is positive semidefinite.

<u>Proof of Proposition 4.2.</u>

Since the objective is continuous in x over a compact set X, it attains its infinum at a point in X.

Q.E.D.

<u>Proposition 4.3.</u> Given the above conditions on $\gamma(t)$, X, and Q, the form $w = Mz + q$, $w, z \geq 0$, $w_i z_i = 0$ has a solution for every $t \in [T_1, T_2]$, and this solution can be found by using the methods of complementary pivot theory in a finite number of pivot operations on M.

<u>Proof of Proposition 4.3.</u>

As stated above, Dantzig and Cottle [3] have shown that the complementary pivot theory algorithm converges to a solution of the quadratic programming problem, if the solution set for $w = Mz + q$, $w, z \geq 0$, is nonempty. Since X is nonempty, and since the infinum of $\gamma(t)'x + x'Qx$ is attained in X for each t, then the necessary and sufficient conditions, i.e., a solution to (4.17), must exist. Thus, the conditions for termination of the algorithm are satisfied.

Q.E.D.

Remembering the characteristics of $\gamma(t)$ explained earlier, we will now show that only a finite number of solutions are considered when solving the parametric programming problem for all t in the finite interval, $[T_1, T_2]$.

Theorem 4.5. There are a finite number of changes in the set of basic variables of (4.17) for the parametric programming problem over a finite interval.

Proof of Theorem 4.5.

Let the M matrix be $(n \times n)$. Since each complementary solution has n basic variables, there are, at most, $\binom{2n}{n}$ complementary solutions. By using the pivoting procedures, the characteristic of any solution is

$$\begin{bmatrix} \bar{w} \\ \\ \bar{z} \end{bmatrix} = \begin{bmatrix} \bar{q}(t) \end{bmatrix} + \begin{bmatrix} \bar{M} \end{bmatrix} \begin{bmatrix} \underline{z} \\ \\ \underline{w} \end{bmatrix} ,$$

where $\bar{w}$ is the set of basic variables of the vector w,

$$w = \begin{bmatrix} \bar{w} \\ \\ \underline{w} \end{bmatrix} ,$$

and $\bar{z}$, the set of basic variables of the vector z. The solution, for any t, is

$$\begin{bmatrix} \bar{w} \\ \\ \bar{z} \end{bmatrix} = \begin{bmatrix} \bar{q}(t) \end{bmatrix} \geq 0 , \qquad \begin{bmatrix} \underline{z} \\ \\ \underline{w} \end{bmatrix} = 0 .$$

The vector function $\bar{q}(t)$ is of the form $f(t)$, and each component of $\bar{q}(t)$ has a finite number of zero crossings in any finite interval of t. Let these points be labelled, $t_i^1, \ldots, t_i^{k_i}$ for the i^{th} complementary solution. The set of points

$$T = \left\{ t_1^1, \ldots, t_1^{k_1}, t_2^1, \ldots, t_2^{k_2}, \ldots, t_{\binom{2n}{n}}^1, \ldots, t_{\binom{2n}{n}}^{k_{\binom{2n}{n}}} \right\}$$

are the only points at which the set of basic variables can change in the finite interval. The set T is countable and has measure zero. If the points are ordered, i.e., $t_0, t_1, \ldots, t_K$, then a set of basic variables that remain basic must exist for any interval $[t_j, t_{j+1})$. This is true because a solution has been shown to exist for every $t \in [t_j, t_{j+1})$, thus the set of basic variables cannot change in this interval.

Q.E.D.

At this point, an algorithm is presented to provide a basic solution to the complementary problem; this solution remains optimal over a positive interval. The method used does not require prior knowledge of the switching points and does not assume nondegenerate solutions at these points. The algorithm employs the same pivoting procedures of the complementary pivot theory algorithm presented in Chapter III.

Definition 4.3. We will say a vector y, which is lexicographically smaller than zero (i.e., $-y \succ 0$) lexico-increases to $\bar{y}$, if $(\bar{y} - y) \succ 0$, and if the component of $\bar{y}$ corresponding to the first nonpositive component of y becomes nonnegative.

Definition 4.4. We will define a lexico-minimum of a set of vectors as that vector y where all the other vectors in it are lexicographically greater than or equal to y.

Definition 4.5. For every t_0, $D_{\bar{q}_i}(t_0)$ is defined as the vector

$$\begin{bmatrix} \bar{q}_i(t_0) \\ \left.\dfrac{d\bar{q}_i(t)}{dt}\right|_{t_0} \\ \vdots \\ \left.\dfrac{d^{N-1}\bar{q}_i(t)}{dt^{N-1}}\right|_{t_0} \end{bmatrix},$$

where $\bar{q}_i(t_0)$ is the current value of the i^{th} component of $\bar{q}(t_0)$ under the pivoting procedures (i.e., those procedures leading to a particular complementary solution). Let

$$\epsilon_i = \inf_{\epsilon} \{\epsilon \mid q_i(t_0 + \epsilon) < 0\} ,$$

and let $\epsilon_0 = \min_i \epsilon_i$.

We will now present an algorithm for solving the parametric quadratic programming problem. It should be noted that this algorithm is also applicable to the parametric linear programming problems discussed in Section IV.A; however, this method is more complicated. A flow chart and proof of the algorithm's finite termination will follow the general iteration given here for any $t = t_0$.

<u>Step I</u>: (a) Solve the complementary pivot theory problem

$$w = Mz + q(t_0)$$

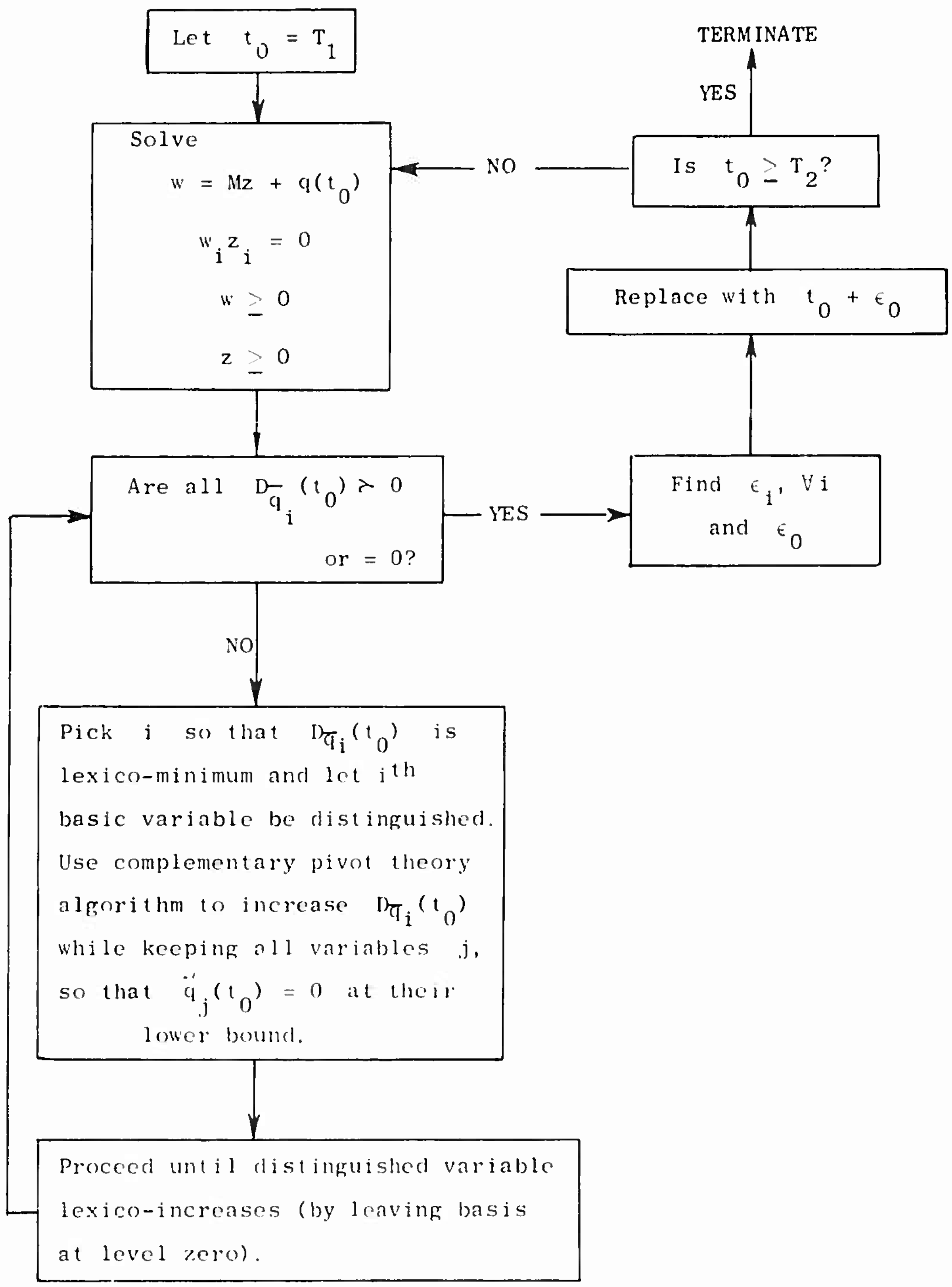
$$w_i z_i = 0, w, z \geq 0 ,$$

for positive semidefinite matrix M, and (b) examine the nonnegative complementary solution for $t = t_0$; i.e., is

$$D_{\bar{q}_i}(t_0) \succ 0 \quad \text{or} \quad = 0 \text{ ?}$$

If $D_{\bar{q}_i}(t_0) \succ 0$ or equal to zero, proceed to Step III; if it is not, go to Step II.

<u>Step II</u>: Choose an index i so that $D_{\bar{q}_i}(t_0)$ is a minimum over all the derivative vectors that are lexicographically less than zero. Retain all other variables of those variables which have derivative vectors lexicographically less than zero, at their present lower bounds, or let them increase (when forced to decrease they are blocking variables). Then let the basic variable corresponding to i be the distinguished variable and



FLOW CHART FOR PARAMETRIC QUADRATIC PROGRAMMING ALGORITHM perform the standard complementary pivot algorithm with the new lower bound restrictions; terminate when the distinguished variable lexicographically increases (i.e., when it drops out of the basis by becoming level 0). The solution of Step I is used to initiate Step II. Now, return to (b) of Step I with the current solution.

<u>Step III</u>: Using the previous definitions, calculate ϵ_i for all i and ϵ_0, the minimum of ϵ_i. The final basis at t_0 is then optimal for $t \in [t_0, t_0 + \epsilon_0]$. If a solution for $t = t_0 + \epsilon_0$ is desired, return to (b) of Step I, using the solution at $t_0 + \epsilon_0$, and proceed with the algorithm.

The procedure given here provides a solution to the parametric quadratic program for any interval of the parameter.

It remains to be shown that each step in the algorithm does terminate in a finite number of executions and that the final basis generated at any t_0 is optimal over a finite positive interval for the quadratic program. Hence the remainder of this section is devoted to this proof.

Step I is solvable in a finite number of steps, if the problem

$$w = Mz + q(t_0)$$

$$w'z = 0, \qquad w, z \geq 0$$

can be solved. Since it is known that the quadratic program has a solution for every t, we are assured of the complementary pivot theory algorithm converging to an optimal solution in a finite number of steps.

<u>Proposition 4.1.</u> Step II must terminate with a complementary solution having the derivative vector of the distinguished variable lexicographically increased while other variables at their lower bound are not lexico-decreased.

<u>Proof of Proposition 1.4.</u>

The complementary pivot theory algorithm when initiated with a basic (complementary) solution, terminates in another complementary solution, since we assume feasibility of the quadratic program. Because the distinguished variable and its complement are converging to a nonnegative complementary solution, this termination occurs when the distinguished variable leaves the basis. When it leaves the basis, the variable has a derivative vector equal to zero; thus it has lexicographically increased. All other variables, at their lower bounds, were not permitted to lexico-decrease; all variables entering the basis are permitted to increase only. Thus the new solution has no variables lexicographically less than zero, and the distinguished variable is lexico-increased.

Q. E. D.

Theorem 4.6. The algorithm terminates in a finite number of steps to a complementary basis that is optimal over a finite positive interval.

Proof of Theorem 4.6.

The execution of Step II lexico-increase at least one of the variables that was lexicographically less than zero while not lexico-decreasing any of them; and it does not introduce any new variables lexicographically less than zero. Since there are only a finite number of lexicographically less than zero vectors (at most, n) and since each has, at most, n components, the execution of Step II must terminate after at most n^2 times with each execution requiring a finite number of steps. The termination condition is that all basic variables are lexicographically greater than or equal to zero (all nonbasic variables lexicographically equal to zero). Thus by Theorem 4.1, the basic variables, and hence the solution to (4.17) are nonnegative over a finite positive interval.

Q. E. D.

From the above results, it has been determined that a finite number of basis changes are required to solve the parametric quadratic programming problem for a finite interval of the parameter. The solution to the parametric quadratic program need not be piecewise constant as it is in the parametric linear programs. For each interval where a single basis remains optimal, the solution will, in fact, have the characteristic form

$$\begin{bmatrix} \bar{w} \\ \bar{z} \end{bmatrix} = \begin{bmatrix} \bar{q}(t) \end{bmatrix} , \qquad w,\underline{z} = 0 ,$$

where $\bar{w}$ and $\bar{z}$ are the complementary variables in their original form, and $\underline{z}$ and $\underline{w}$ are their complements. $\bar{q}(t)$ is nonnegative over the interval and has the form where each of its components solves some particular N^{th} order homogeneous constant coefficient linear differential equation. When the value of some $\bar{q}_i(t)$ goes negative for some t, the basic variables $\bar{w},\bar{z}$ are no longer optimal, and a new set of basic variables must be found with the complementary property. The new values of $\bar{q}_i(t)$ are just linear combinations of the former components.

Chapter V

GENERALIZED PROGRAMMING ALGORITHM FOR OPTIMAL CONTROL PROBLEMS

The mathematical programming results obtained in the previous two chapters are applied to the linear system, continuous-time optimal control problems to formulate [9] a generalized linear program. A solution procedure based on this formulation is then developed and is shown to terminate in an optimal solution, i.e., an optimal control is provided to the continuous-time problem.

A. Formulation

The control problems will now be formulated as generalized programs and then the subproblems will be shown to be parametric programming problems of the form presented in Chapter IV. The control problem can be restated as

$$\min_{u(\cdot)} J = \int_0^T \dot{x}_0(t)\ dt = x_0(T) \tag{5.1a}$$

$$\dot{x}(t) = Fx(t) + Gu(t) \tag{5.1b}$$

$$x \in E^n , \qquad u \in E^m , \qquad \text{and}$$

$$\bar{x} = \begin{bmatrix} x_0 \\ x \end{bmatrix} \in E^{n+1} ,$$

$$x(0) \in S_0 , \qquad \text{and} \qquad x(T) \in S_T ,$$

$$u(t) \in U = \{u | Au \geq b\} , \qquad U \subseteq E^m ,$$

F is an $n \times n$ real matrix, and

G is an $n \times m$ real matrix.

$$\dot{x}_0(t) = f_0'x(t) + g_0'u(t) + h \sum_{i=1}^{m} |u_i(t)| + u(t)'Qu(t) \qquad (5.1c)$$

where f_0 is a fixed real n-vector and g_0 is a fixed real m-vector, h is a real constant, and Q is an $n \times n$ real matrix.

We will first consider strictly linear cost functionals, i.e., $h = 0$, $Q = 0$.

Letting

$$\bar{F} = \left[\begin{array}{c|c} 0 & f_0' \\ \hline \vdots & \\ 0 & F \end{array}\right], \quad \text{and}$$

$$\bar{G} = \left[\begin{array}{c} g_0' \\ \hline G \end{array}\right], \qquad (5.2)$$

the completely linear system can be expressed by

$$\dot{\bar{x}}(t) = \bar{F}\bar{x}(t) + \bar{G}u(t) . \qquad (5.3)$$

When a particular vector function $u^i(t)$ and an initial condition $\bar{x}^i(0)$ are given, the solution to (5.3), at $t = T$, is

$$\bar{x}(T) = e^{T\bar{F}} x^i(0) + \int_0^T e^{(T-t)\bar{F}} \bar{G}u^i(t)\ dt , \qquad (5.4a)$$

and the solution to (5.1b) is

$$x(T) = e^{TF} x^i(0) + \int_0^T e^{(T-t)F} Gu^i(t)\ dt . \qquad (5.4b)$$

The set $S_0^F \subseteq E^m$ is defined

$$S_0^F \equiv \{y | x \in S_0, y = e^{TF} x\} .$$

If S_0 is a convex set, S_0^F is also a convex set by the linear mapping. The set $\mathcal{S} \subseteq E^n$ can be defined by

$$\mathcal{S} \equiv S_T - S_0^F , \qquad \text{or}$$

$$\mathcal{S} = \left\{ z | z = x - y, \ x \in S_T, \ y \in S_0^F \right\} .$$

<u>Proposition 5.1.</u> If S_T and S_0^F are convex, then $\mathcal{S}$ is convex.

<u>Proof of Proposition 5.1.</u>

Let x^1, x^2 be points in S_0^F and let y^1, y^2 be points in S_T; then $x^1 - y^1$, and $x^2 - y^2$ are in $\mathcal{S}$. For all λ, $0 \leq \lambda \leq 1$,

$$\lambda(x^1-y^1) + (1-\lambda)(x^2-y^2) = \lambda x^1 + (1-\lambda)\ x^2 - \lambda y^1 - (1-\lambda)\ y^2 .$$

$$\lambda x^1 + (1-\lambda)\ x^2 \in S_0^F \qquad \text{since } S_0^F \text{ is convex.}$$

$$\lambda y^1 + (1-\lambda)\ y^2 \in S_T \qquad \text{since } S_T \text{ is convex.}$$

Thus $\lambda(x^1-y^1) + (1-\lambda)(x^2-y^2) \in \mathcal{S}$, and implies that $\mathcal{S}$ is convex.

Q.E.D.

Let $\bar{S}_0$ and $\bar{S}_T$ be defined as

$$\bar{S}_0 = \left\{ \bar{y} \in E^{n+1} \middle| \bar{y} = \begin{bmatrix} y_0 \\ y \end{bmatrix} , \quad y_0 = 0, y \in S_0 \right\}$$

and

$$\bar{S}_T = \left\{ \bar{y} \in E^{n+1} \middle| \bar{y} = \begin{bmatrix} y_0 \\ y \end{bmatrix} , \quad y_0 = 0, y \in S_T \right\};$$

the sets $\bar{S}_0^F$ and $\bar{\mathcal{S}}$ are similarly defined.

Using the above set of definitions, we can restate the initial and final state constraints of the fixed time control problems as

$$x(0) = 0 \, . \qquad x_0(0) = 0 \, , \qquad \text{and}$$

$$x(T) \in \mathcal{S} \, .$$

Thus, it can be assumed, without loss of generality, that the system initiates at the origin with no prior costs.

1. Control Problems Formulated as Generalized Programs

If we take the vector functionals of the control $P = P[u(t)]$ to be defined by

$$P = \int_0^T e^{(T-t)F} \, Gu(t) \, dt \, , \tag{5.5a}$$

and

$\overline{P} = \overline{P}[u(t)]$ to be defined by

$$\overline{P} = \int_0^T e^{(T-t)\overline{F}} \, \overline{G}u(t) \, dt \, , \tag{5.5b}$$

then, let

$$C \equiv \left\{ P \,\middle|\, u(t) \in U \, , \quad P = \int_0^T e^{(T-t)F} \, Gu(t) \, dt \right\}$$

and

$$\overline{C} \equiv \left\{ \overline{P} \,\middle|\, u(t) \in U \, , \quad \overline{P} = \int_0^T e^{(T-t)\overline{F}} \, \overline{G}u(t) \, dt \right\} \, .$$

Proposition 5.2. If U is a convex set, the set $C(\overline{C})$ is convex.

Proof of Proposition 5.2.

Let $u^1(t)$, $u^2(t)$ be vector functions in U, for all t, and

$$P^1 = \int_0^T e^{(T-t)F} Gu^1(t)\, dt \quad \in C$$

$$P^2 = \int_0^T e^{(T-t)F} Gu^2(t)\, dt \quad \in C\ .$$

For all λ, $0 \leq \lambda \leq 1$, $\bar{\lambda} = (1-\lambda)$,

$$\lambda P^1 + \bar{\lambda} P^2 = \lambda \int_0^T e^{(T-t)F} Gu^1(t)\, dt + \bar{\lambda} \int_0^T e^{(T-t)F} Gu^2(t)\, dt$$

$$= \int_0^T e^{(T-t)F} G[\lambda u^1(t) + \bar{\lambda} u^2(t)]\, dt\ .$$

Since U is convex, $\lambda u^1(t) + \bar{\lambda} u^2(t) \in U$, $\forall t$.
Thus $\lambda P^1 + \bar{\lambda} P^2 \in C$. and C is convex.

Q.E.D.

Remembering that the state at time 0 is assumed to be at the origin and using the definitions of P, C, and Eq. (5.4b), we find C is equivalent to R_T, the reachable set of U at time T. The control problem is feasible, iff

$$C \cap \mathcal{S} \neq \emptyset\ .$$

Given a specified control function $u(t)$, the cost associated with that control is $J[u(t)]$, since the cost at time $t = 0$ is zero. Note that

$$\bar{P} = \begin{bmatrix} p_0 \\ P \end{bmatrix}, \qquad \text{where} \quad p_0 = J[u(t)]\ ,$$

when $u(t)$ is the control generating P by Eq. (5.5a). Thus

$$\overline{C} = \left\{ \overline{P} \in E^{n+1} \middle| \overline{P} = \begin{bmatrix} p_0 \\ p \end{bmatrix} , \quad P = P(u) , \quad p_0 = J(u) , \quad u(t) \in U \right\} .$$

Note also that by using Proposition 5.2, $\overline{C}$ is convex.

Let $U_0' = (1,0,\ldots,0)$ and note that the first component of the vectors $\overline{S}$, in the set $\overline{\mathcal{S}}$, is defined to be zero. Also note that the first component of the $\overline{P}$ vector represents the cost of using the control (and its corresponding trajectory) generating $\overline{P}$. Thus we are looking for a vector $\overline{P} \in \overline{C}$, a vector function $u(t)$ generating $\overline{P}$, and a vector $\overline{S} \in \overline{\mathcal{S}}$ to satisfy

$$\max_{\overline{P} \in \overline{C}} \lambda , \qquad \mu, \nu \geq 0$$

$$\text{subject to} \qquad U_0\lambda + \overline{P}\mu = \overline{S}\nu$$

$$\mu = 1$$

$$\nu = 1 , \tag{5.6}$$

where μ and ν are scalars. Maximizing λ is equivalent to minimizing $J[u(t)]$, the first component of the vector $\overline{P}$, where $u(t)$ generates $\overline{P}$. Since $\overline{P}$ must be taken from a convex set $\overline{C}$ and $\overline{S}$ must be taken from a convex set $\overline{\mathcal{S}}$, the above formulation is a generalized program of the Dantzig-Wolfe type. In the following chapter we will show that an optimal solution to the control problem is an optimal solution to the generalized programming problem. We will now show that any solution to the generalized programming problem is an optimal solution to the original control problem.

A solution to the generalized programming problem consists of a vector P in the reachable set R_T, a control function $u(t)$ in the admissible control region U generating P, and a vector S in the constraint set of terminal states $\mathcal{S}$, so that

$$P \equiv S .$$

The above equality insures the transformation of the system from an initial point, $x(0) \in S_0$ to a final point $x(T) \in S_T$ by the vector function $u(t)$, chosen from U. Thus it is a feasible control. By minimizing J over all feasible sets of $\overline{P}$ and $\overline{S}$, we can find a feasible solution with the least cost. This is precisely an optimal solution to the continuous-time control problem.

2. Generalized Programming Subproblems as Parametric Programs

To complete the generalized programming formulation, its subproblem must be described. Here we assume there are at least $n + 2$ vectors P^i and/or S^i available to provide a feasible solution to (5.6), so that the problem

$$\max_{\mu,\nu} \lambda , \qquad \mu,\nu \geq 0$$

$$U_0\lambda + P^1\mu_1 + P^2\mu_2 + \ldots + P^j\mu_j = S^{j+1}\nu_1 + \ldots + S^{j+p}\nu_p$$

$$\mu_1 + \mu_2 + \ldots + \mu_j = 1$$

$$\nu_1 + \nu_2 + \ldots + \nu_p = 1 , \tag{5.7}$$

is solvable and has a dual solution vector

$$\overline{\pi}' = (\pi_0, \pi', \pi_{n+1}, \pi_{n+2}) ,$$

where $\pi' = (\pi_1, \ldots, \pi_n)$. The subproblem is then formulated in two parts:

$$\min_{\overline{S} \in \overline{\mathcal{S}}} \overline{\pi}' \begin{bmatrix} \overline{S} \\ 0 \\ 1 \end{bmatrix} , \qquad \text{and} \tag{5.8}$$

$$\min_{\overline{P} \in \overline{C}} \overline{\pi}' \begin{bmatrix} \overline{P} \\ 1 \\ 0 \end{bmatrix} . \tag{5.9}$$

The solution to (5.8) is dependent on the explicit definition of the set $\mathcal{S}$ or $\bar{\mathcal{S}}$, the simplest case being the fixed end point problem, which consists of a single element. In this case, the subproblem (5.8) is trivial and need not be considered. If $\bar{\mathcal{S}}$ is a convex polyhedral set, then (5.8) is a linear program that needs to be solved once for each iteration of the master problem:

Subproblem (5.9) can be described

$$\min_{\bar{P} \in \bar{C}} \bar{\pi}' \begin{bmatrix} \bar{P} \\ 1 \\ 0 \end{bmatrix} = \min_{\bar{P} \in \bar{C}} \bar{\pi}' \begin{bmatrix} \int_0^T e^{(T-t)\bar{F}} \bar{G}u(t)\, dt \\ 1 \\ 0 \end{bmatrix} . \qquad (5.10)$$

Since the requirement $\bar{P} \in \bar{C}$ is equivalent to the requirement $u(t) \in U$ for all t, (5.10) becomes

$$\min_{u(t) \in U} \bar{\pi}' \begin{bmatrix} \int_0^T e^{(T-t)\bar{F}} \bar{G}u(t)\, dt \\ 1 \\ 0 \end{bmatrix} , \quad \text{or}$$

since $\bar{\pi}$ does not depend on t,

$$\min_{u(t) \in U} \left[\int_0^T (\pi_0, \pi')\, e^{(T-t)\bar{F}} \bar{G}u(t)\, dt \right] + \pi_{n+1} . \qquad (5.11)$$

The minimum of the integral is attained when the integrand is minimized at every point. Let

$$\gamma(t) = (\pi_0, \pi')\, e^{(T-t)\bar{F}} \bar{G} \qquad (5.12)$$

be an m-dimensional vector function. Thus the subproblem becomes

$$\text{Find} \quad u(t) \in U , \qquad t \in [0,T]$$

$$\text{so that} \quad \gamma(t)'u(t) \quad \text{is a minimum .} \tag{5.13}$$

From (5.12), it is obvious that $\gamma(t)$ has the property of each of its components being a member of the class of solutions to an $n + 1^{st}$ order, homogenous, constant coefficient, linear differential equation. Since our attention is restricted to those U that are polyhedral sets, (5.13) becomes

$$\min \gamma(t)'u(t)$$

$$Au(t) \geq b$$

$$t \in [0,T] . \tag{5.14}$$

(Note that the inequality may be reversed or an equality without loss of generality.) Thus a solution $u(t)$, for the subproblem, can be obtained by using the parametric linear programming methods of Chapter IV.

In a similar manner, we can formulate minimum fuel, minimum time, and quadratic loss in control problems as generalized programs. It can also be shown that the minimum fuel and minimum time are special cases of the linear loss problems just described. Since generalized programming can be applied to general convex programming problems, we can formulate optimal control problems with loss functions convex in the control variable as generalized programs. However, only the quadratic loss in the control case will be discussed in detail, since this is (to the author's knowledge) the only general nonlinear convex loss function which has a known finite solution procedure for the parametric subproblem. Separable piecewise linear (convex) functions of the control can also be formulated as a special case of the linear loss problem, although it will not be shown here.

The previous formulations can be generalized if we observe the following general linear system control problem and use the notation given in Chapter II,

$$\min_{u(t) \in U} \quad J = \int_0^T x_0(t)\ dt,$$

$$\text{where} \quad \bar{x}(t) = \begin{bmatrix} x_0(t) \\ x(t) \end{bmatrix} \in E^{n+1} , \quad \text{and}$$

$$\dot{\bar{x}}(t) = \overline{F}\bar{x}(t) + \bar{G}u(t) + f(u)\ U_0 , \qquad (5.15)$$

$$\text{where} \quad f(u) = \begin{cases} 0 & , \text{ linear loss} \\ \sum_i |u_i| & , \text{ minimum fuel} \\ u'Qu & , \text{ quadratic loss ,} \end{cases}$$

where Q is positive semidefinite, $u(t) \in U$, $t \in [0,T]$, $x(0) \in S_0$, $x(T) \in S_T$, and S_0, S_T are convex sets in E^n. We note that J is a convex functional in $u(t)$, since $u(t)$ is a vector sequence drawn from a convex set U, and $f(u)$ is a convex function in u. Thus the solution to Eq. (5.15) can be noted as

$$\bar{x}(T) = e^{\overline{F}T}\ \bar{x}(0) + \int_0^T e^{\overline{F}(T-t)}\ \bar{G}u(t)\ dt + \int_0^T e^{\overline{F}(T-t)}\ f[u(t)]\ U_0\ dt . \qquad (5.16)$$

Now, we define P and $\bar{P}$, as before, by

$$P = \int_0^T e^{F(T-t)}\ Gu(t)\ dt , \quad \text{and} \qquad (5.17)$$

$$\bar{P} = \begin{bmatrix} p_0 \\ p \end{bmatrix} = \int_0^T e^{\overline{F}(T-t)}\ \bar{G}u(t)\ dt + \int_0^T f[u(t)]\ dt\ U_0 , \qquad (5.18)$$

where U_0 is an $n+1$-dimensional unit vector with a one in the first component. Thus the second integral of (5.18) becomes

$$\int_0^T \sum_{i=1}^{m} |u_i(t)|\,dt \,, \qquad \text{minimum fuel} \,,$$

$$\int_0^T u(t)'Qu(t)\,dt \,, \qquad \text{quadratic loss} \,.$$

Also as was done before, we can now define $\overline{C}$ as

$$\overline{C} = \left\{ \overline{P} \in E^{n+1} \;\middle|\; \overline{P} = \begin{bmatrix} z \\ \int_0^T e^{F(T-t)}\, Gu(t)\, dt \end{bmatrix} , \right.$$

$$\left. z \geq J[u(t)] \,, \quad u(t) \in U, \quad \forall t \right\} .$$

Thus the $\overline{P}$, as defined in (5.18), are members of $\overline{C}$. We also note that $\overline{C}$ is a convex set, since $J[u(t)]$ is a convex functional.

The vectors $\overline{S}$ and set $\overline{\mathcal{S}}$ are defined as before. Thus the general linear system control problem can be formulated as a generalized programming problem,

$$\text{Find } \overline{P} \in \overline{C} \,, \quad \overline{S} \in \overline{\mathcal{S}} \,, \quad \text{to}$$

$$\max_{\overline{P} \in \overline{C}} \lambda \,, \qquad \mu, \nu \geq 0$$

$$\text{subject to} \quad U_0\lambda + \overline{P}\mu = \overline{S}\nu$$

$$\mu = 1$$

$$\nu = 1 \,. \tag{5.19}$$

Again we note that the solution to the generalized program is one in which a vector $\overline{P}^*$ is found so that the last n elements P^* belong to the set $\mathcal{S}$, and, out of all possible vectors, $P \in \mathcal{S}$, i.e., feasible solutions to the control problem; the first component of $\bar{P}^*$, taken to be the value of the loss functional, is minimal. Thus the vector control function $u^*(t)$ generating $\bar{P}^*$ is a solution to the optimal control problem.

We now describe the subproblem corresponding to the minimum fuel and quadratic loss problems as in Eqs. (5.8) and (5.9). For the general case, following the similar reasoning given for the linear loss case, the subproblem to the restricted master problem becomes, (remembering that $\pi_0 = 1$, since $(U_0', 0, 0)$ is a basis vector and

$$\bar{\pi}' \begin{bmatrix} U_0 \\ 0 \\ 0 \end{bmatrix} = 1 \;)$$

$$\min_{u(\cdot)} \int_0^T \left\{ f[u(t)] + (\pi_0 \pi')' e^{(T-t)\bar{F}} \bar{G}u(t) \right\} dt + \pi_{n+1} \; . \qquad (5.20)$$

Defining $\gamma(t)$ as before, (5.20) becomes

$$\min_{u(t)} \gamma(t)'u(t) + f[u(t)]$$

$$\text{subject to } u(t) \in U \; , \quad \forall t \; . \qquad (5.21)$$

For the quadratic loss problem, with $U = \{u \mid Au \geq b, \; u \geq 0\}$, (5.21) becomes a parametric quadratic programming problem in $u(t)$ of the form discussed in Chapter IV [due to the form of $\gamma(t)$].

We now look at the minimum fuel problem for two classes of U. The first class has the classical form of the minimal fuel problem, where

$$U = \{u \mid |u_i(t)| \leq 1, \quad i = 1, \ldots, m\} \; ,$$

and the second class is a general polyhedral U.

For the first case, (5.21) has the following solution (the singular arcs, $\gamma(t) \equiv 1$ for an interval, are not discussed because no solution is defined),

$$
\begin{aligned}
-1 < \gamma_i(t) < 1 \quad &, \qquad u_i(t) = 0 \\
\gamma_i(t) < -1 \quad &, \qquad u_i(t) = 1 \\
\gamma_i(t) > 1 \quad &, \qquad u_i(t) = -1 \\
&\text{for all } i \, . \qquad\qquad (5.22)
\end{aligned}
$$

The magnitude of one as a bound for the control is noted to be nonrestrictive, since G can be scaled to permit other values. The formulation can also be adapted to treat lower bounds on $u_i(t)$ with magnitudes that differ from the upper bounds. These changes affect the ranges of $\gamma_i(t)$ in (5.22). Thus the subproblem for the standard minimum fuel problem has a well-defined solution, and its execution in relation to the master problem is proportionately as quick, regardless of the size of the control space.

The minimum fuel problem for general polyhedral sets U has a subproblem equivalent to that of the linear loss case. Following the same steps given above, the subproblem for general U which replaces (5.23) is

$$\min_{u(\cdot)} \; \gamma(t)u(t) + \sum_i |u_i(t)|$$

subject to

$$
\begin{aligned}
Au(t) &\geq b \\
0 \leq t &\leq T \, . \qquad\qquad (5.23)
\end{aligned}
$$

This is equivalent to a parametric linear programming problem of the type presented in Chapter IV when the variable $u_i(t)$, which is unrestricted in sign, is replaced by the difference between two nonnegative variables,

$$u_i(t) = \bar{u}_i(t) - \bar{\bar{u}}_i(t) \; . \tag{5.24}$$

The constraints are replaced by

$$\begin{aligned} A\bar{u} - A\bar{\bar{u}} &\geq b \\ \bar{u} &\geq 0 \\ \bar{\bar{u}} &\geq 0 \; . \end{aligned} \tag{5.25}$$

Since linear programming algorithms consider only basic solutions, $\bar{u}_i$ and $\bar{\bar{u}}_i$ cannot be basic at the same time because their columns A_i and $-A_i$, respectively, are linearly dependent. Thus for every i, either $\bar{u}_i$ ‹ ‑ $\bar{\bar{u}}_i$ must be at level zero. Using this result, $|u_i(t)|$ can be replaced by

$$|u_i(t)| = \bar{u}_i(t) + \bar{\bar{u}}_i(t) \; , \tag{5.26}$$

and the equivalent parametric linear program is

$$\min_{\bar{u},\bar{\bar{u}}} \sum_i [\gamma_i(t) + 1] \, u_i(t) + \sum_i [-\gamma_i(t) + 1] \, \bar{\bar{u}}_i(t)$$

$$\begin{aligned} \text{subject to} \qquad A\bar{u}(t) - A\bar{\bar{u}}(t) &\geq b \\ \bar{u}_i(t), \; \bar{\bar{u}}_i(t) &\geq 0 \\ \text{for} \quad 0 \leq t \leq T \; . \end{aligned} \tag{5.21}$$

Since a generalized programming formulation is shown to be used for finding a feasible solution to the linear system control problems, the minimal time problem can be solved with these methods. The subproblem to the generalized programming problem for feasible solutions is also shown to be a parametric linear programming problem of the type discussed. Here we will present a solution procedure (similar to the one proposed by Ho [11]) to the minimal time problem that is equivalent to finding, if it exists, a feasible solution to the generalized programming formulation of linear systems and, of course, to determining whether or not a feasible solution exists.

The minimal time problem can be stated as

$$\min_{u(\cdot)} \int_0^{t_f} dt = t_f$$

$$(t_f \text{ is free}),$$

$$\text{where } x(0) \text{ is given and } x(t_f) = 0, \text{ and}$$

$$\dot{x}(t) = Fx(t) + Gu(t)$$

$$x(t) \in E^n, \quad u(t) \in U \subseteq E^m . \tag{5.28}$$

Thus we wish to find the shortest time in which we can transfer the state from a given point to the origin or, equivalently, to find the minimum time for which a control $u(t) \in U$ exists to transfer the system from a given point to the origin. The latter statement is the one related to a generalized programming formulation of the minimal time problem. Define P as before, for any fixed T, and $S = -e^{TF} x(0)$. Let the set of all P be $C \equiv R_T$, i.e., the reachable set, which is convex. We want to know whether

$$S \in R_T , \qquad \text{for any fixed } T , \tag{5.29}$$

so that we can find the minimum T for which (5.29) holds. Equivalently, we want to know whether a solution exists for

$$P\mu = S$$

$$\mu = 1$$

$$P \in C \equiv R_T , \qquad \text{for any } T , \tag{5.30}$$

and again we want to find the minimum T for which a solution to (5.30) exists. The solution procedure for (5.30) for any T, will be discussed later. At this point, we will solve the minimum time problem by choosing a T, and try to solve (5.30). If a solution exists, decrease T and proceed; if one does not exist then increase T and proceed. If the increments for the increase and decrease of T are chosen wisely, this procedure will converge to an answer to the minimum time problem.

The solution to (5.30) is a phase I generalized programming procedure that is also used to find initial feasible solutions to the optimal control problems discussed here. A finite convergence procedure is shown for phase I methods when its existence is known, and a test for its existence will be presented for control problems for which the existence of feasible solutions is not assumed.

To generalize and summarize the above results, the following class of control problems may be formulated as generalized programming problems:

$$\min_{u(\cdot)} \quad J = \int_0^T f(x,u)\, dt \ ,$$

$$x = Fx + Gu,$$

$$x(0) \in S_o \ , \quad x(T) \in S_T \ , \quad u(t) \in U \ , \quad 0 \leq t \leq T \ , \qquad (5.31)$$

where S_0, S_T, and U are convex, and

$$f(x,u) = f_1(x) + f_2(u)$$

where $f_1(x)$ is linear in x and independent of u, $f_2(u)$ is convex in u and independent of x. When S_0, S_T, and U are polyhedral sets and $f_2(u)$ is quadratic in u or the sum of the absolute value of the components (with linear terms permitted), the generalized programming problem is solvable by the methods presented in the previous two chapters. The rest of this chapter is devoted to the development of the algorithm for solving these generalized programming problems and to pointing out the specific features of the algorithm so that it can be adopted for special purposes, including the determination of feasibility or its existence.

B. Solution of the Control Problem

The first step and, at times, the major problem in the solution of the control problem is to find an admissible control yielding a feasible solution. An important characteristic of the generalized programming solution of optimal control problems is that, at every stage in the optimization phase, a feasible solution is always available. With this feasible solution, a bound on the value of the optimal objective function can then be computed. Thus, if the solution procedure is interrupted before its convergence to an optimal solution, a feasible solution can be recovered and an estimate of how close it is to an optimal solution provided. This estimate or bound can be used to terminate the algorithm, since suboptimal solutions having an objective value close to the optimal one, are generally sufficient for decision purposes. Although the general solution to the linear control problems may be an infinite convergent process, the generation of a feasible solution, if interior solutions exist, is a finite process, and the generation of a suboptimal solution, as close as desired in objective value to the optimal, can also be obtained in a finite number of steps. The algorithm and its variants are presented in this section along with convergence and finiteness proofs. The characteristics of solutions and their relations to known results in control theory are presented in Chapter VI.

1. Generation of a Feasible Solution

There are two major aspects of finding a feasible solution. The first of these is the determination of whether or not a feasible solution exists; the second is to generate the feasible solution, if it does exist. In both cases, a phase I procedure of the generalized programming problem is used. The control problem to be considered is

$$\text{Find a } u(t) \in U = \{u \in E^m \mid Au \leq b\}, \qquad \forall t,$$

such that $x(0) \in S_0$ and

$x(T) \in S_T$, when x is controlled by

$$\dot{x}(t) = Fx(t) + Gu(t) . \tag{5.32}$$

Without loss of generality, we let $S_0 = 0$ and $S_T = \mathcal{S}$.

Let us assume that the reachable set R_T is a continuum. Let us also assume that, if

$$R_T \cap \mathcal{S} \neq \emptyset ,$$

then $R_T \cap \mathcal{S}$ is a continuum. This condition insures the existence of a finite-dimensional neighborhood in that set of desired final states in the reachable set. As will be shown, these conditions imply that the phase I portion of the generalized programming formulation of the control problem terminates with a feasible solution in a finite number of of steps.

Since the convergence (Chapter III) of the generalized program assumes that a nondegenerate feasible starting solution is available, the phase I procedure must terminate with such a nondegenerate feasible solution. This implies, for an n-dimensional state space, that a collection of $n + 1$ vectors P must be generated so that the initial basis,

$$B = \begin{bmatrix} P^1 & P^2 & \cdots & P^{n+1} \\ \\ 1 & 1 & \cdots & 1 \end{bmatrix},$$

is nonsingular. Also a value, $S \in \mathcal{S}$, must be provided so that $B^{-1}S$ is a vector and is strictly positive in all components.

Although convergence proofs require the results of the following phase I procedure, more efficient variants (to be presented) should be used in practice and can provide a feasible solution in fewer steps for most problems.

We will now show the procedure for finding a feasible solution when $\mathcal{S}$ is a single point and a ball of radius ρ in E^n (a ρ-neighborhood) is also contained in the reachable set, R_T. The ρ-neighborhood is used to avoid degeneracy problems, in much the same manner as a simplex lexicographic methods by perturbating the original right-hand side. Thus we seek to generate a set of $n + 1$ vectors P^i, to provide a nondegenerate solution to the set of linear equations.

$$P^1\mu_1 + \dots + P^{n+1}\mu_{n+1} = S$$

$$\mu_1 + \dots + \mu_{n+1} = 1$$

$$\mu_i > 0 \ , \ \forall i \ . \tag{5.33}$$

a. Some Properties of Convex Sets

Definition 5.1. The convex hull of a set, $X \subseteq E^n$, is the intersection of all convex sets in E^n containing X.

Definition 5.2. The convex hull $\triangle$ of a finite set of $n + 1$ points, $x_1, x_2, \dots, x_{n+1}$, in E^n is an n-dimensional simplex, if the flat of minimal dimension containing $\triangle$ has dimension n. The points x_i are called vertices.

Lemma 5.1 [12]. If $\triangle$ is an n-dimensional simplex with vertices $x_i (i = 1, \dots, n + 1)$, then $\triangle$ consists of all points $x \in E^n$ for which constants α_i exist, so that

$$x = \sum_{i=1}^{n+1} \alpha_i x_i \ , \qquad \sum \alpha_i = 1$$

$$\alpha_i \geq 0 \ .$$

Definition 5.3. A set of $k+1$ points in E^n is geometrically independent, if no $(k-1)$-dimensional hyperplane contains all the points.

Definition 5.4. A set $\{x_0, x_1, \ldots, x_n\}$ of vectors in E^n is pointwise independent (algebraic counterpart of geometrically independent), if the k vectors, $x_1 - x_0$, $x_2 - x_0$, $\ldots$, $x_k - x_0$ are linearly independent.

Lemma 5.2 [13]. If $X = \{x_0, x_1, \ldots, x_k\}$ is a pointwise independent set in E^n, then there exists a unique k-dimensional hyperplane H^k containing X having the property that $x \in H^k$, iff

$$x = \sum_{i=0}^{k} \alpha_i x_i \ , \qquad \sum_{i=0}^{k} \alpha_i = 1 \ ,$$

where the α_i are unique. The α_i are the barycentric coordinates of x with respect to X.

Let us look at the convex hull of the set of $n+1$ points in E^n, X,

$$X = \{x_0, k_1, \ldots, x_n\} \ , \qquad \text{where}$$

$$x_0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad x_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \ \ldots, \quad x_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

and denote it by convX. These points are pointwise independent, thus their convex hull is a simplex.

Let

$$x^* = \begin{bmatrix} \frac{1}{n+1} \\ \vdots \\ \frac{1}{n+1} \end{bmatrix},$$

thus $x^* \epsilon$ convX, since the barycentric coordinates are $\alpha_i = 1/n+1$, $\forall i$. Also since $\alpha_i > 0$, $\forall i$, x^* is in the interior of convX.

Define a new set of points,

$$X' = \{x'_0, x'_1, \ldots, x'_n\} ,$$

$$\text{where} \quad x'_i = x_i + \Delta x_i ,$$

$$x_i \quad \text{is as before, and}$$

$$\|\Delta x_i\| < \frac{1}{(n+1)^2} .$$

Lemma 5.3. The points x'_i are pointwise independent.

Proof of Lemma 5.3.

Let $x''_j = x_j + \Delta x_j - \Delta x_0$, $j = 1, \ldots, n$. Assume that the x''_j are linearly dependent. Then a nontrivial set of λ_j exists, so that

$$\sum_{j=1}^{n} \lambda_j x''_j = 0 .$$

Thus,

$$\sum_j \lambda_j x''_j = \sum_j \lambda_j [x_j + \Delta x_j - \Delta x_\eta] = 0 ,$$

and

$$\sum_j \lambda_j[\Delta x_0 - \Delta x_j] = \sum_j \lambda_j x_j \ . \tag{5.34}$$

Taking the vector inner product of both sides of (5.34) with x_k,

$$\left(\sum_j \lambda_j x_j\right)' x_k = \lambda_k \ , \qquad \text{and}$$

$$\sum_j \lambda_j(\Delta x_0 - \Delta x_j)' x_k < \sum_j \lambda_j\left[\frac{2}{(n+1)^2}\right] \ ,$$

since

$$\Delta x_j' x_k \leq \|\Delta x_j\| \cdot \|x_k\| < \frac{1}{(n+1)^2} \ ,$$

Thus we have

$$\sum_j \lambda_j\left[\frac{2}{(n+1)^2}\right] > \lambda_k \ . \tag{5.35}$$

If we take inner products with all the x_j, $j = 1, \ldots, n$, and sum the left- and right-hand sides of the result (5.35), we obtain

$$\sum_k \lambda_k < \sum_k \left[\sum \lambda_j \frac{2}{(n+1)^2}\right] = \frac{2}{(n+1)^2} \sum_j \lambda_j \ ,$$

which implies $1 < 2n/(n+1)^2$, which is a contradiction for $n \geq 0$. Thus the vectors x'_j are pointwise independent or geometrically independent.

Q.E.D.

<u>Theorem 5.1.</u> The point x^* is in the interior of convX'.

<u>Proof of Theorem 5.1.</u>

ConvX' forms a simplex, since the points x'_i are pointwise independent and form an n-dimensional hyperplane. Thus

$$x^* = \sum_{i=0}^{n} \alpha'_i x'_i \, , \qquad \text{and}$$

$$\sum_i \alpha'_i = 1 \tag{5.36}$$

has a unique solution in the α'_i.

We must now show that $\alpha'_i > 0$, $\forall i$. Notice that

$$\sum_{i=0}^{n} \frac{1}{n+1} x_i = x^* = \sum \alpha'_i x'_i \, ,$$

or

$$\sum_{i=0}^{n} \left[\alpha'_i x'_i - \frac{1}{n+1} x_i \right] = 0 \, . \tag{5.37}$$

Since $x'_i = x_i + \Delta x_i$, (5.37) becomes

$$\sum_{i=0}^{n} \left[\left(\alpha'_i - \frac{1}{n+1} \right) x_i + \alpha'_i \Delta x_i \right] = 0 \, . \tag{5.38}$$

By taking the inner product of (5.38) with each x_i, $i = 1, \ldots, n$ successfully, we get the set of equations,

$$\alpha'_1 - \frac{1}{n+1} + \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i, x_1 \rangle = 0$$

$$\alpha'_2 - \frac{1}{n+1} + \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i, x_2 \rangle = 0$$

$$\vdots$$

$$\alpha'_n - \frac{1}{n+1} + \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i, x_n \rangle = 0 ,$$

or, in general

$$\alpha'_j = \frac{1}{n+1} - \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i, x_j \rangle, \ j = 1, \ldots, n .$$

Since

$$\langle \Delta x'_i, x_i \rangle < \frac{1}{(n+1)^2} ,$$

$$\alpha'_i > \frac{1}{n+1} - \frac{1}{(n+1)^2} \sum_{i=0}^{n} \alpha'_i = \frac{1}{(n+1)} - \frac{1}{(n+1)^2} .$$

Thus $\alpha'_i > 0$ for $i = 1, \ldots, n$. Now,

$$\sum_{i=0}^{n} x_i = \underline{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} ,$$

and we take the vector product of (5.38) with $\underline{1}$ to get

$$\sum_{i=1}^{n}\left(\alpha'_i - \frac{1}{n+1}\right) + \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i,\ \underline{1} \rangle = 0\ ,$$

which implies by using

$$\sum_{i=1}^{n} \alpha'_i = 1 - \alpha'_0\ , \qquad \text{and}$$

$$\sum_{i=1}^{n} \frac{1}{n+1} = 1 - \frac{1}{n+1}\ ,$$

that

$$1 - \alpha'_0 - 1 + \frac{1}{n+1} + \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i,\ \underline{1} \rangle = 0\ , \qquad \text{or}$$

$$\alpha'_0 = \frac{1}{n+1} + \sum_{i=0}^{n} \alpha'_i \langle \Delta x'_i,\ \underline{1} \rangle\ .$$

Remembering that $\|\Delta x_i\| < 1/(n+1)^2$ and $\|\underline{1}\| = \sqrt{n}$, we obtain

$$\alpha'_0 > \frac{1}{n+1} - \frac{\sqrt{n}}{(n+1)^2} \sum_{i=0}^{n} \alpha'_i = \frac{1}{n+1} - \frac{\sqrt{n}}{(n+1)^2}\ .$$

Thus $\alpha'_0 > 0$, and x^* is in the interior of convX'.

Q. E. D.

By using the assumption that a . -neighborhood about $S = X_T$ is also in R_T, the following set of points are found in R_T,

$$X_\rho = \left\{ X_T - \frac{1}{n+1}\left(\frac{\rho}{n+1}\right) \underline{1} + \frac{\rho}{n+1} e_i, \ i = 0, 1, \ldots, n \right\},$$

where e_i, $i = 1, \ldots, n$ is an n-dimensional unit vector with a one in the i^{th} row and e_0 is the null vector. Geometrically, X_ρ defines the vertices of an n-dimensional simplex with X_T defined by the barycentric coordinates of $1/(n+1)$, for each point in X_ρ.

$$\text{Let} \quad X_i = X_T - \frac{1}{n+1}\left(\frac{\rho}{n+1}\right) \underline{1} + \frac{\rho}{n+1} e_i, \qquad \forall i .$$

<u>Proposition 5.6.</u> Let a ball of radius $[\rho/(n+1)][1/(n+1)]^2$, about X_i for any i, be N_i, then

$$N_i \subseteq R_T .$$

<u>Proof of Proposition 5.6.</u>

$$X_i - X_T = -\frac{1}{n+1}\left(\frac{\rho}{n+1}\right) \underline{1} + \frac{\rho}{n+1} e_i$$

$$\|X_i - X_T\| \leq \left\|\frac{1}{n+1}\left(\frac{\rho}{n+1}\right) \underline{1}\right\| + \left\|\frac{\rho}{n+1} e_i\right\|$$

$$\|X_i - X_T\| \leq \frac{1}{n+1}\left(\frac{\rho}{n+1}\right)\sqrt{n} + \frac{\rho}{n+1} = \frac{\rho}{n+1}\left(\frac{\sqrt{n}+n+1}{n+1}\right).$$

The maximum distance from any point in N_i to X_i is $\frac{\rho}{n+1}\frac{1}{n+1}^2$, thus the maximum distance from any point in N_i to X_T is d,

$$d \leq \frac{\rho}{n+1}\left(\frac{1}{n+1}\right)^2 + \frac{1}{(n+1)^2}[\rho][\sqrt{n}+n+1], \quad \text{or}$$

$$d \leq \rho\left[\frac{1}{(n+1)^2}\left\{\left(\frac{n+2}{n+1}\right) + n + \sqrt{n}\right\}\right],$$

which is less than ρ for $n \geq 1$. Thus $N_i \subseteq R_T$, since a ρ-neighborhood about X_T is in R_T.

Q.E.D.

We will now solve the following problems, for all $X_i \in X_\rho$:

b. <u>Solution Procedure for Feasible Solution to a Generalized Program</u>

Find a vector P (or a convex combination of vectors), so that

$$w < \frac{\rho}{n+1}\left(\frac{1}{n+1}\right)^2$$

$$w = \sum_{i=1}^{n} y_i^+ + \sum_{i=1}^{n} y_i^-$$

$$P\mu + Iy^+ - Iy^- = X_i$$

$$\mu = 1$$

$$\mu, y_i^+, y_i^- \geq 0 \,, \qquad P \in C \,. \tag{5.39}$$

This is the phase I procedure of a generalized program. The columns P^i can now be generated to the master program (a linear program),

$$\min_{\mu, y} w = \sum_{i=1}^{n} y_i^+ + \sum_{i=1}^{n} y_i^-$$

$$P^1\mu_1 + \ldots + P^k\mu_k + Iy^+ - Iy^- = X_i$$

$$\mu_1 + \ldots + \mu_k \qquad = 1$$

$$\mu_i, y_i^+, y_i^- \geq 0 \,. \tag{5.40}$$

From the vector of dual variables $\bar{\pi}^k$, the k^{th} iteration of the generalized program (5.40) generates a new column P^{k+1}, by means of the subproblem,

$$\min_{P \in C} \bar{\pi}^{k\,'} \begin{bmatrix} P \\ 1 \end{bmatrix} . \tag{5.41}$$

This subproblem is equivalent to finding a vector function $u^{k+1}(t)$ that

$$\text{minimizes} \quad \pi^{k'} e^{F(T-t)} Gu(t)$$

$$\text{subject to} \quad u(t) \in U ,$$

$$\text{where} \quad \bar{\pi}^{k'} = \left(\pi^{k'}, \pi^k_{n+1}\right) . \tag{5.42}$$

The problem (5.42) for polyhedral U has previously been shown to be a parametric linear programming problem and is solvable by the methods introduced in Chapter IV.

The generation of P^{k+1} from $u^{k+1}(t)$ is

$$P^{k+1} = \int_0^T e^{F(T-t)} Gu^{k+1}(t)\, dt .$$

The generalized program (5.39) terminates when

$$w^k < \frac{\rho}{(n+1)} \left(\frac{1}{n+1}\right)^2$$

for some iteration k of (5.40). We know that the minimum value of w for all $P \in C$ is zero, since $X_i \in R_T$, and

$$P^{i*} = X_i \in C .$$

Therefore

$$P^{i*}\mu + Iy^{+} - Iy^{-} = X_i$$

$$\mu \qquad = 1$$

μ, y_i^+, $y_i^- \geq 0$, has a solution with all y_i^+, $y_i^- = 0$. The terminating condition for each part of the phase I procedure terminates a generalized program at a suboptimal solution with the objective value some specified distance from its optimal value. Thus the generalized program terminates in a finite number of steps.

Once the value of w becomes low enough, the solution to the phase I procedure for each X_i must be recovered. From the final solution, for each X_i phase, let

$$u^{i*}(t) = \sum_{i=1}^{k} u^i(t)\mu_i \qquad \text{and}$$

$$P^{i*} = \sum_{i=1}^{k} P^i\mu_i$$

for each $i = 0, 1, \ldots, n$. Also note that

$$P^{i*} = \sum_{i=1}^{k} P^i\mu_i = \sum_{i=1}^{k} \int_0^T e^{F(T-t)}\, G[u^i(t)]\, dt\, \mu_i$$

$$= \int_0^T e^{F(T-t)}\, Gu^{i*}(t)\, dt\ .$$

<u>Proposition 5.7.</u> The set of vectors P^{i*} and controls $u^{i*}(t)$ constitute a nondegenerate feasible solution to the set of equations,

$$P^{0*}\mu_0 + P^{1*}\mu_1 + \dots + P^{n*}\mu_n = X_T$$

$$\mu_0 + \mu_1 + \dots + \mu_n = 1$$

$$\mu_i > 0 \ , \quad \forall i \ . \tag{5.43}$$

Proof of Proposition 5.7.

The points X_i can be transformed and scaled to the points.

$$x_0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad x_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad x_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

and X_T can be scaled to

$$\sum_i \left(\frac{1}{n+1}\right) x_i = x^* \ .$$

By the same transformation, P^{i*} can be replaced by

$$x_i' = x_i + \Delta x_i \ ,$$

where

$$\|\Delta x_i\| < \left(\frac{1}{n+1}\right)^2 \ .$$

Thus the results of Theorem 5.1 hold and

$$x^* = \sum \alpha_1' x_i$$

$$\sum \alpha_i = 1$$

$$\alpha_i > 0 \ , \quad \forall i \ ,$$

and by retransforming and rescaling, $\mu_i = \alpha_i' > 0$, $\forall i$ for (5.43).

Q.E.D.

c. Application to Control Problems

The above procedure may be used to initiate the optimum control problem when feasibility is known. However, in cases where the existence of a feasible solution is not known, another phase I procedure can be used. This procedure will determine feasibility and in the process provide an initial feasible solution.

For the fixed end point problem, a solution with $w = 0$ to the generalized program,

$$\min_{\mu, y} \quad w = \sum_{i=1}^{n} y_i^+ + \sum_{i=1}^{n} y_i^-$$

$$P\mu + Iy^+ + Iy^- = S$$

$$\mu = 1$$

$$y_i^+, \ y_i^- \geq 0$$

$$P \in C \ , \tag{5.44}$$

implies that a vector $P \in C$ exists which provides a feasible solution to the control problem. If the optimal solution to (5.44) has a value of $w^* > 0$, a feasible solution does not exist to the control problem.

Theorem 5.2. If at any stage k in the solution of (5.44), the value of $w^k + \delta^k > 0$, the original control problem is infeasible, where δ^k is the objective value of the subproblem to the generalized program (5.44).

Proof of Theorem 5.2.

Consider the master program at the k^{th} stage, as: $\max_{\mu} \lambda$

$$\text{subject to } U_0\lambda + \bar{P}^1\mu_1 + \ldots + \bar{P}^k\mu_k + \bar{I}y^+ - \bar{\bar{I}}y^- - \bar{S}\nu = 0$$

$$\mu_1 + \ldots + \mu_k = 1$$

$$y_i^+,\ y_i^-,\ \mu_i \geq 0 , \tag{5.45}$$

where

$$\bar{P}^i = \begin{bmatrix} 0 \\ \\ p^i \end{bmatrix} , \qquad \bar{S} = \begin{bmatrix} 0 \\ \\ S \end{bmatrix} ,$$

$$\bar{I} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ & & I & \end{bmatrix} , \quad \bar{\bar{I}} = \begin{bmatrix} -1 & -1 & \ldots & -1 \\ & & I & \end{bmatrix} .$$

Thus $\lambda^k = -w^k$ and the dual variable to (5.45) is

$$\bar{\pi}^{k'} = \left(\pi_0^k,\ \pi^{k'},\ \pi_{n+1}^k,\ \pi_{n+2}^k\right) .$$

By using the duality theorem of linear programming and the dual to (5.45),

$$\lambda^k = \pi_{n+1}^k + \pi_{n+2}^k .$$

Also

$$\delta^k = \min_{P \in C} \pi^{k'} \begin{bmatrix} 0 \\ P \\ 1 \\ 0 \end{bmatrix} = \min_{P \in C} \pi^{k'} P + \pi_{n+1}^k .$$

Since $\bar{S}$ is always in the basis of (5.45),

$$\bar{\pi}^{k'} \begin{bmatrix} -\bar{S} \\ 0 \\ 1 \end{bmatrix} = 0 \quad \text{implies}$$

$$\pi^{k'}(-S) = -\pi^{k}_{n+2} \, . \tag{5.46}$$

By hypothesis,

$$0 < \delta^{k} + w^{k} = \min_{P \in C} \pi^{k'} P + \pi^{k}_{n+1} + w^{k} \, , \qquad \text{or}$$

$$0 < \delta^{k} + w^{k} \leq \pi^{k'} P + \pi^{k}_{n+1} + w^{k}$$

for any $P \in C$.

Since $w^{k} = -\lambda^{k} = -\pi^{k}_{n+1} - \pi^{k}_{n+2}$,

$$0 < \delta^{k} + w^{k} \leq \pi^{k} P - \pi^{k}_{n+2} \, , \qquad \text{and by (5.46)}$$

$$0 < \delta^{k} + w^{k} \leq \pi^{k'} P - \pi^{k'} S \, , \qquad \text{or}$$

$$0 < \delta^{k} + w^{k} \leq \pi^{k'} (P-S) \, , \qquad \text{for all } P \, .$$

Thus, for any $P \in C$,

$$0 < \pi^{k'} (P-S) \, ,$$

which implies that

$$\pi^{k} \neq 0 \qquad \text{and}$$

$$P - S \neq 0 \qquad \text{for all } P.$$

Thus there is no admissible control function $u(t)$ which can generate a feasible solution to the control problem.

Q.E.D.

However, if at any iteration k, $w^k = 0$, the current solution is a feasible solution to the control problem and phase II (the optimization phase) of the generalized program can be initiated. Since this phase I procedure, unlike the previous one, is not necessarily a finite process, the optimization phase may begin when $w^k \leq \epsilon$, some small positive number and the desired final point not precisely attained. For any practical control problem, when feasibility is not known, a point at some arbitrarily small distance away from a determined fixed point would be an acceptable terminal point for the control problem. Thus, the phase I procedure would be finite even when feasibility is not assumed.

For the variable end point problem, phase I procedures are much simpler. For example, when the desired final region $\mathcal{S}$ is constrained to lie in some r-neighborhood about a determined point S, the phase I procedures outlined previously are used $(r < \rho)$. If feasibility is assumed, a series of reachable points in the r-neighborhood should be chosen, and a phase I procedure identical to the first one discussed in this paper would provide a nondegenerate feasible solution. If a nondegenerate solution is not necessary, then a procedure identical to that for finding the existence of a feasible solution can be used. In this case S is allowed to be the right-hand side of (5.44) and the algorithm is terminated when $w \leq r$, as the following theorem points out.

Theorem 5.3. If $w^k \leq r$, then the solution P^* to (5.44) at stage k, which produces the value w^k, is a vector which lies in an r-neighborhood about S, and it is a feasible solution for the control problem.

Proof of Theorem 5.3.

Since $w \leq r$,

$$\sum_{i=1}^{n} y_i^+ + \sum_{i=1}^{n} y_i^- = \sum_{i=1}^{n} |(P_i^* - S_i)| \leq r \,,$$

thus

$$\sum_{i=1}^{n} + \sqrt{(P_i^* - S_i)^2} \leq r, \qquad \text{and}$$

$$\|P^* - S\| = + \sqrt{\sum_i (P_i^* - S_i)^2} \leq \sum_i + \sqrt{(P_i^* - S_i)^2} \leq r \,.$$

Q.E.D.

A general phase I procedure for the variable end point problem can be used to determine feasibility as well as to find a feasible solution when it exists. We consider the problem

$$\min w = \sum_{i=1}^{n} y_i^+ + \sum_{i=1}^{n} y_i^-$$

subject to

$$P + Iy^+ - Iy^- + S = 0$$

$$= 0$$

$$= 1$$

$$P \in C$$

$$S \in \mathcal{S} \,, \qquad \text{or equivalently,} \tag{5.47}$$

a restricted master problem at iteration k,

$$\max \lambda$$

$$U_0\lambda + \sum_{i=1}^{k} \overline{P}^i \mu_i - \sum_{j=1}^{k} \overline{S}^j \nu_j + \overline{I}y^+ - \overline{\overline{I}}y^- = 0$$

$$\sum_{i=1}^{k} \mu_i \qquad\qquad = 1$$

$$\sum_{j=1}^{k} \nu_j \qquad\qquad = 1$$

$$\mu_i,\ \nu_i,\ y_i^+,\ y_i^- \geq 0\ , \qquad \text{and } \overline{P},\ \overline{S},\ \overline{I},\ \overline{\overline{I}} \text{ are as before .} \tag{5.48}$$

When a solution to (5.48) produces a value of $\lambda^k = 0$ or equivalently $w^k = 0$, the solution is a feasible solution to the control problem. The following two theorems will prove useful for determining when a feasible solution does not exist and for recovering a feasible solution (when it does exist) from a particular iteration of the master problem (5.48). Let the subproblem of (5.48) be designated by the parameters, δ^k and Δ^k, and defined by

$$\delta^k = \min_{P \in C} \overline{\pi}^{k'} \begin{bmatrix} 0 \\ P \\ 1 \\ 0 \end{bmatrix}, \qquad \text{and} \tag{5.49}$$

$$\Delta^k = \min_{S \in \mathcal{S}} \overline{\pi}^{k'} \begin{bmatrix} 0 \\ S \\ 0 \\ 1 \end{bmatrix}, \tag{5.50}$$

where $\overline{\pi}^k$ is the dual variable to (5.48).

Theorem 5.4. If at any stage k of the initial phase procedure for the variable end point problem $w^k + \delta^k + \Delta^k > 0$, the original control problem is infeasible.

Proof of Theorem 5.4.

Letting $\bar{\pi}^{k'} = \left(\pi_0^k, \pi^{k'}, \pi_{n+1}^k, \pi_{n+2}^k\right)$ and remembering from the duality theorem, that

$$w^k = -\lambda^k = -\pi_{n+1}^k - \pi_{n+2}^k ,$$

then, by hypothesis,

$$0 < \delta^k + \Delta^k + w^k = \min_{P \in C} \bar{\pi}^{k'} \begin{bmatrix} 0 \\ P \\ 1 \\ 0 \end{bmatrix} + \min_{S \in \mathcal{A}} \bar{\pi}^{k'} \begin{bmatrix} 0 \\ -S \\ 0 \\ 1 \end{bmatrix} + w^k .$$

Substituting in the above,

$$0 < \delta^k + \Delta^k + w^k \leq \bar{\pi}^{k'} \begin{bmatrix} 0 \\ P \\ 1 \\ 0 \end{bmatrix} + \bar{\pi}^{k'} \begin{bmatrix} 0 \\ -S \\ 0 \\ 1 \end{bmatrix} - \pi_{n+1}^k - \pi_{n+2}^k$$

for all $P \in C$ and $S \in \mathcal{A}$.

Thus

$$0 < \delta^k + \Delta^k + w^k \leq \pi^{k'} P + \pi^{k'} (-S) + \pi_{n+1}^k + \pi_{n+2}^k - \pi_{n+1}^k - \pi_{n+2}^k , \quad \text{or}$$

$$0 < \pi^{k'} (P-S) ,$$

which implies that

$$\pi^k \neq 0 \qquad \text{and}$$

$$(P-S) \neq 0 \qquad \text{for all} \quad P \in C \quad \text{and} \quad S \in \mathcal{S}\,.$$

Therefore, there is no feasible solution to the control problem.

Q.E.D.

Theorem 5.4 indicates an infeasibility condition, and in the next theorem (Theorem 5.5) a feasibility condition is presented.

<u>Theorem 5.5.</u> If at any iteration k the solution to (5.48) provides a vector,

$$S^* = \sum_{j=1}^{k} S^i \nu^k_i$$

and a value $w^k \leq r$ so that S^* is an interior point in $\mathcal{S}$ and has an r-neighborhood surrounding S that is also in $\mathcal{S}$, the original control problem is feasible and the solution

$$P^* = \sum_{i=1}^{k} P^i \nu^k_i \qquad [\nu^k_1 \text{ being a solution to } (5.48)]$$

is a feasible, reachable point.

<u>Proof of Theorem 5.5.</u>

Since $w^k \leq r$, the vector,

$$P^* = \sum_{i=1}^{k} P^i \nu^k_i$$

has the property

$$\|P^* - S^*\| = +\sqrt{\sum_i (P_I^* - S_i^*)^2} \leq \sum_i +\sqrt{(P_i^* - S_i^*)^2} = w^k \leq r \quad .$$

Thus the point P^* is in an r-neighborhood about S^* implying that $P^* \in \mathcal{S}$. Hence the control generating P^* is a feasible control.

Q.E.D.

2. Generation of an Optimal Solution

For the fixed end point problem, the subproblem for phase I methods is a parametric linear program of the type discussed in Chapter IV. For some variable end point problems, the phase I procedure has an added subproblem,

$$\min_{S \in \mathcal{S}} \pi^{k'} \begin{bmatrix} 0 \\ -S \\ 0 \\ 1 \end{bmatrix} , \quad \text{or}$$

$$\min_{S \in \mathcal{S}} - \pi^{k'} S \tag{5.51}$$

Thus we seek a vector S in a specified set $\mathcal{S}$ that minimizes the sum

$$\sum \pi_i^k (-s_i) \ ,$$

subject to the constraint that $S \in \mathcal{S}$. If $\mathcal{S}$ is a convex polyhedral set, this problem is a linear program that must be solved once for each iteration of the master problem. For other classes of $\mathcal{S}$, the subproblem depends on the definition of $\mathcal{S}$.

While discussing the optimization phase of the generalized programming formulation of the optimal control problem, we will consider only the fixed end point problem.

To avoid degeneracy, it is desirable to start the optimization phase with a set of $n + 1$ vectors P^i that provide a nonsingular basis,

$$B = \begin{bmatrix} U_0 & \bar{P}^1 & \bar{P}^2 & \cdots & \bar{P}^{n+1} \\ 0 & 1 & 1 & & 1 \end{bmatrix},$$

which is feasible for the program,

$$\begin{aligned} &\max_{\mu} \lambda \\ &\text{subject to} \quad U_0\lambda + \bar{P}^1\mu_1 + \ldots + \bar{P}^{n+1}\mu_{n+1} = \bar{S} \\ &\qquad\qquad\qquad \mu_1 + \ldots + \mu_{n+1} = 1 \\ &\qquad\qquad\qquad \mu_i \geq 0 , \end{aligned} \qquad (5.52)$$

where $$\bar{P}^i = \begin{bmatrix} \int_0^T f(x, u^i)\, dt \\ P^i \end{bmatrix},$$

where $u^i(t) \in U$ generates the vector, P^i, and

$$\bar{S} = \begin{bmatrix} 0 \\ S \end{bmatrix}.$$

These vectors are immediately available when using the phase I procedure discussed initially. For other phase I procedures, once a feasible solution is found and assuming a neighborhood about that solution is also feasible, the initial phase I procedure may be used by generating the right-hand sides in a similar manner about the known feasible solution.

Restating the fixed end point optimal control problem as

$$\dot{x} = Fx + Gu$$

$$x(0) = 0, \; x(T) = S$$

$$x \in E^n \; , \quad u(t) \in U \subseteq E^m \; , \quad \forall t \; ,$$

where $$U = \{u \,|\, Au \geq b\} \; ,$$

and

$$\min_{u(\cdot)} \; J = \int_0^T f(x,u) \; dt \; ,$$

where

$$f(x,u) = \begin{cases} f_0'x + g_0'u \\ f_0'x + g_0'u + \sum_{i=1}^{m} |u_i| \\ f_0'x + g_0'u + u'Qu \end{cases}$$

we can define $\bar{F}$ and $\bar{G}$, as before.

When a set of vector control functions $u^i(t)$ is given from the phase I procedure, the vectors $\bar{P}^i$ must be generated.

$$\bar{P}^i = \begin{bmatrix} \int_0^T (f_0'x + g_0'u) \; dt + \int_0^T f(u) \; dt \\ \\ p \end{bmatrix} \; ,$$

or

$$\bar{P}^i = \int_0^T e^{\bar{F}(T-t)} \; \bar{G}u(t) \; dt + \int_0^T f(u) \; dt \; p_0 \; .$$

Since the matrix e^{Ft} must be provided for the phase I method, the matrix,

$e^{\bar{F}t}$ is easily shown to be

$$
e^{\bar{F}t} = \left[\begin{array}{c:c} 0 & f_0' e^{Ft} \\ \hdashline 0 & \\ \vdots & \\ 0 & e^{Ft} \end{array}\right]
$$

by considering the system

$$
x_0(t) = f_0' x \qquad x \in E^n
$$

$$
x\,(t) = Fx
$$

where $x(t) = e^{F(t-\tau)}\, x(\tau)$, and

$$
x_0(t) = f_0' x(t) = f_0' e^{F(t-\tau)}\, x(\tau) .
$$

Letting $\bar{x} = \begin{bmatrix} x_0 \\ x \end{bmatrix}$,

$$
\bar{x}(t) = e^{\bar{F}(t-\tau)}\, \bar{x}(\tau) = \left[\begin{array}{c:c} 0 & f_0' e^{F(t-\tau)} \\ \hdashline 0 & \\ \vdots & e^{F(t-\tau)} \\ 0 & \end{array}\right] \begin{bmatrix} x_0(\tau) \\ x(\tau) \end{bmatrix} .
$$

From (5.52) we see that the linear equations in [illegible]$_i$ always insure that a feasible solution exists for the control problem. The simplex method, when applied to the master problem, maintains primal feasibility at all times even when augmenting the linear equations with a vector $\bar{P}^{k+1}$. Thus at any time in the execution of the optimization phase, a feasible solution is available from the current basis. Also as will be shown, a bound on the optimal solution is provided at each stage.

In the following let $J^k = -\lambda^k$, where λ^k is the objective value of the master problem in the k^{th} iteration.

Theorem 5.6. During any k^{th} iteration of the optimization phase of the generalized program, the optimal value of the cost function $J(u^*)$ satisfies the following inequalities [2]:

$$J^k(\hat{u}) + \delta^k \leq J(u^*) \leq J^k(\hat{u}) \leq J^k = -\lambda^k \,, \tag{5.53}$$

where

$$\hat{u}(t) = \sum_{i=1}^{k} \mu^i(t) \,.$$

Proof of Theorem 5.6.

Consider the equivalent linear program, as before,

$$\begin{aligned}
&\max_{\mu} \lambda \\
&U_0\lambda + \bar{P}^1\mu_1 + \bar{P}^2\mu_2 + \ldots + \bar{P}^k\mu_k - S\nu = 0 \\
&\mu_1 + \mu_2 + \ldots + \mu_k = 1 \\
&\nu = 1 \\
&\mu_i \geq 0 \,.
\end{aligned} \tag{5.54}$$

The solution to (5.54) is λ^k and by the dual theorem,

$$\lambda^k = \pi^k_{n+1} + \pi^k_{n+2} \,,$$

where $\bar{\pi}^{k'} = \left(\pi^k_0, \pi^{k'}, \pi^k_{n+1}, \pi^k_{n+2}\right)$ is the dual variable to (5.54). Since $\pi^k_0 = 1$, the subproblem has the solution,

$$\delta^k = \min_{u(t) \in U} \left\{ J(u) + \pi^{k'} P + \pi^k_{n+1} \right\} .$$

Thus for the value u^*,

$$\delta^k \leq J(u^*) + \pi^{k'} P^* + \pi^k_{n+1} , \quad \text{and}$$

$$-\lambda^k + \delta^k = J^k + \delta^k \leq J(u^*) + \pi^{k'} P^* + \pi^k_{n+1} - \pi^k_{n+1} - \pi^k_{n+2} ,$$

or

$$J^k + \delta^k \leq J(u^*) + \pi^{k'} P^* - \pi^k_{n+2} .$$

Since

$$\bar{\pi}^{k'} \begin{bmatrix} 0 \\ -S \\ 0 \\ 1 \end{bmatrix} = -\pi^{k'} S + \pi^k_{n+2} = 0 ,$$

$$J^k + \delta^k \leq J(u^*) + \pi^{k'} (P^* - S) .$$

If u^* is the optimal u, then P^* is feasible and $P^* = S$. Thus,

$$J^k + \delta^k \leq J(u^*) \leq J^k(\hat{u}), \qquad \text{where} \quad J^k(\hat{u})$$

is the current solution and $\hat{u}$ is a feasible control, and the right-hand inequality follows immediately since $J(u^*)$ is the minimum cost.

Since $J(u)$ is convex in u,

$$J^k(\hat{u}) \leq J^k = \sum_{i=1}^{k} \mu_i J(u^i) .$$

Therefore

$$J^k(\hat{u}) + \delta^k \leq J^k + \delta^k \leq J(u^*) \leq J^k(\hat{u}) \leq J^k \tag{5.55}$$

Q.E.D.

Corollary 5.1. When $\delta^k = 0$, $J^k(\hat{u}) = J^k = J(u^*)$.

Proof of Corollary 5.1.

The proof follows from (5.55), since equality holds throughout.

Q.E.D.

Note that the value, λ^k, of the master problem is an approximation to the current solution at iteration k. At any stage, the solution defined by

$$\hat{u}(t) = \sum_{i=1}^{k} \mu_i^k u^i(t)$$

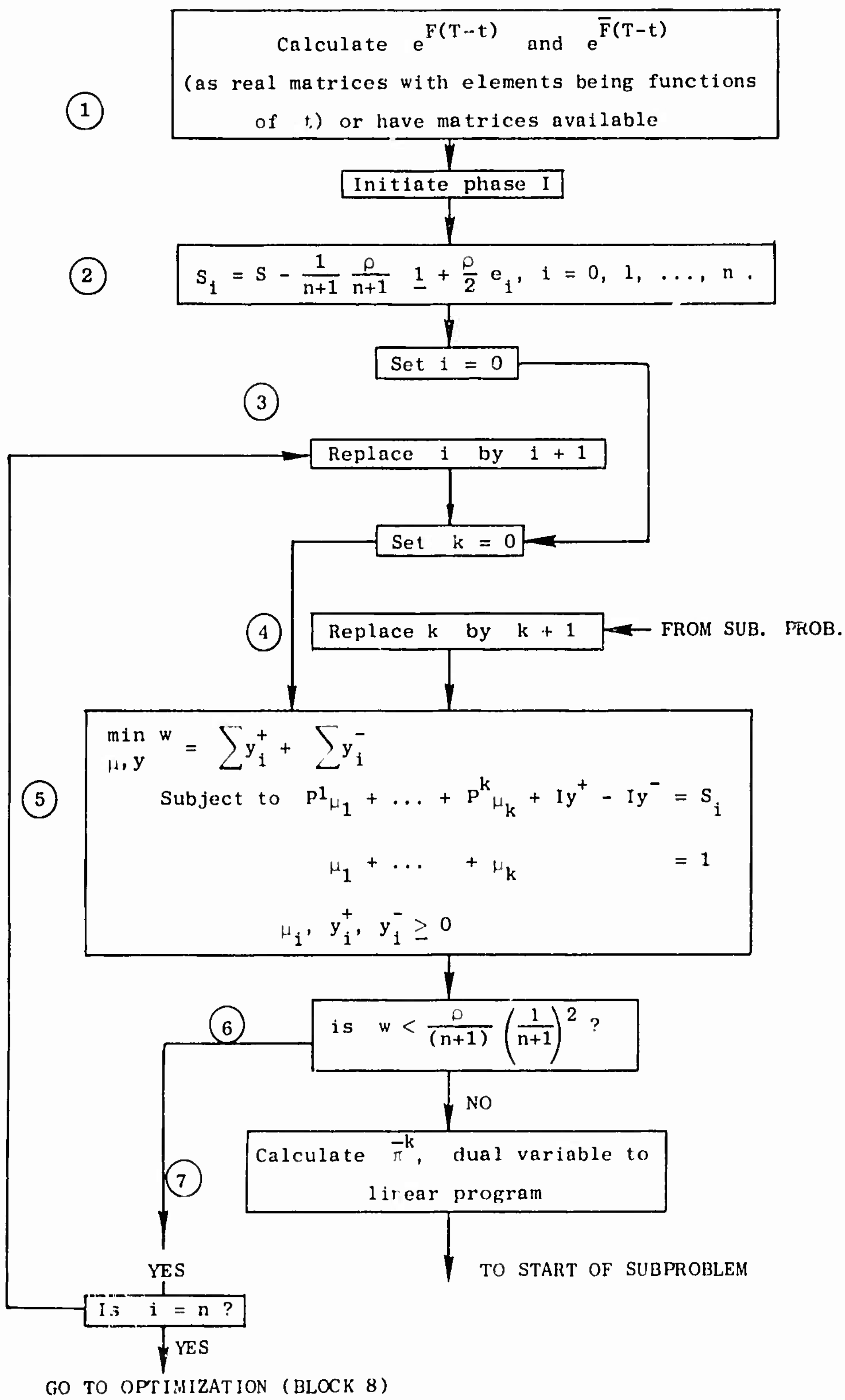
has a cost $J(\hat{u})$ which by convexity is smaller than J^k.

Although δ^k does not necessarily increase monotonically to 0, it does so for a subsequence of k. Thus the best bound from previous iterations should be retained until a better bound is attained. The current value (at iteration k of the generalized program) of δ^k may be used to provide a stopping condition for termination of the optimization phase. By observing the value

$$\delta^k/J^k ,$$

we can determine the maximum percentage by which the objective function can decrease for the optimal solution, and we are assured that the current solution is feasible to the original control problem.

We now present a flow chart of the generalized programming solution to the optimal control problem. We will use the fixed end point problem for an example, assume a ρ-neighborhood about S is in R_T, and apply the long form of the phase I procedure.



GENERALIZED PROGRAMMING ALGORITHM FOR LINEAR OPTIMAL CONTROL.

OPTIMIZATION

8

Let $P^i = \sum_j P^j \mu^i_j$ and

$u^i(t) = \sum_j u^j \mu^i_j$ from i^{th} problem

for all $i = 0, 1, \ldots, n$

phase II

Set $k = n$

9

Calculate p^i_0 from $u^i(t)$ and $e^{\overline{F}(T-t)}$

for $i = 0, 1, \ldots, n$

10

Replace k by $k + 1$ ← FROM SUB

11

$\max_\mu \lambda$

Subject to

$$\lambda + p^0_0\mu_0 + p^1_0\mu_1 + \ldots + p^k_0\mu_k = 0$$

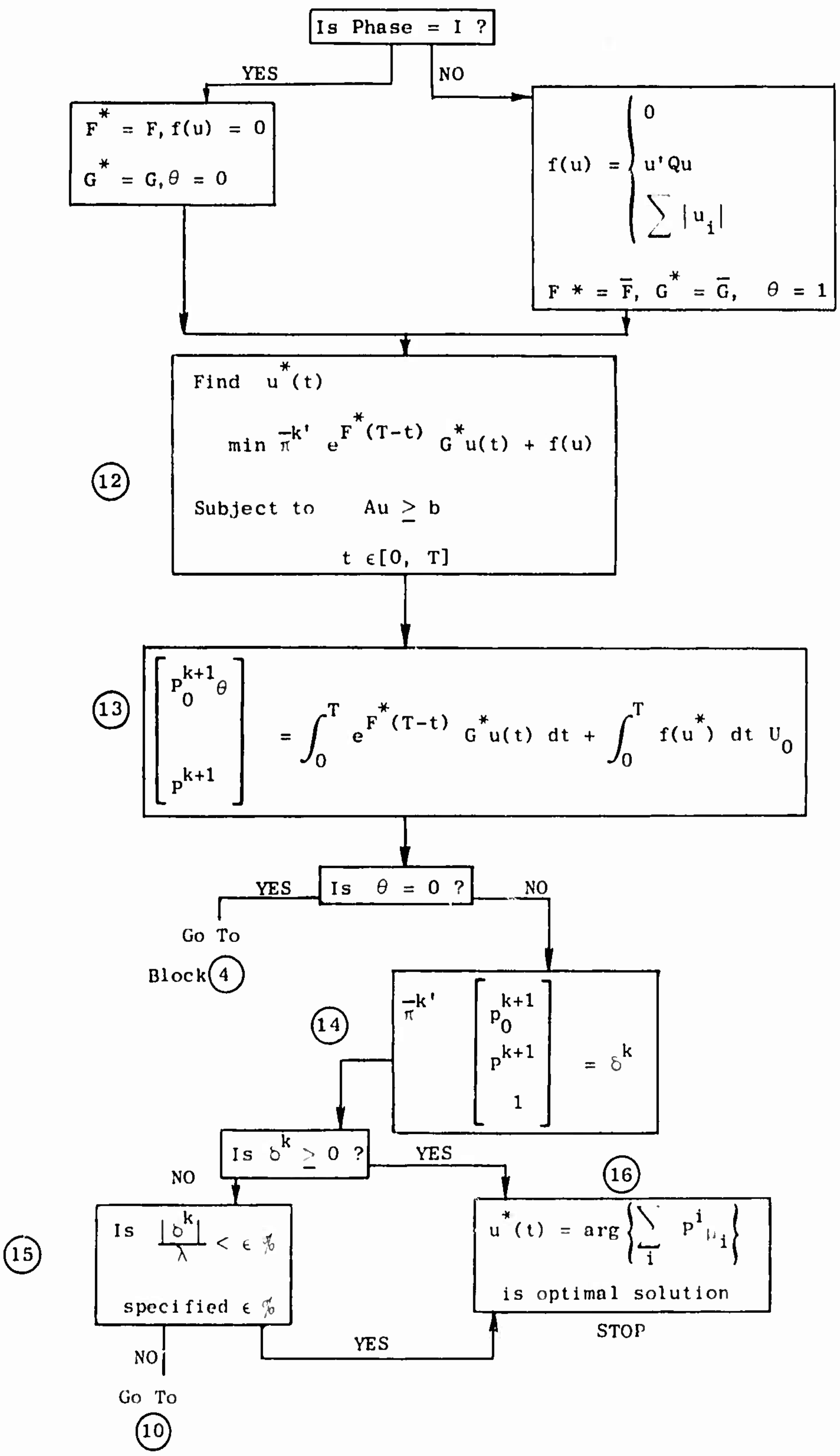
$$P^0\mu_0 + P^1\mu_1 + \ldots + P^k\mu_k = S$$

$$\mu_0 + \mu_1 + \ldots + \mu_k = 1$$

$$\mu_i \geq 0 \quad i = 0, 1, \ldots, k$$

Calculate $\bar{\pi}^k$ - dual to linear program

TO SUBPROBLEM


Is Phase = I ?
YES
NO
$F^* = F, f(u) = 0$
$G^* = G, \theta = 0$
$f(u) = \begin{cases} 0 \\ u'Qu \\ \sum |u_i| \end{cases}$
$F^* = \bar{F}, G^* = \bar{G}, \theta = 1$
12
Find $u^*(t)$
$\min \bar{\pi}^{k'} e^{F^*(T-t)} G^* u(t) + f(u)$
Subject to $Au \geq b$
$t \in [0, T]$
13
$\begin{bmatrix} p_0^{k+1} \theta \\ p^{k+1} \end{bmatrix} = \int_0^T e^{F^*(T-t)} G^* u(t)\, dt + \int_0^T f(u^*)\, dt\, U_0$
Is $\theta = 0$?
YES
NO
Go To Block 4
14
$\bar{\pi}^{k'} \begin{bmatrix} p_0^{k+1} \\ p^{k+1} \\ 1 \end{bmatrix} = \delta^k$
Is $\delta^k \geq 0$?
NO
YES
15
Is $\frac{|\delta^k|}{\lambda} < \epsilon \%$
specified $\epsilon \%$
16
$u^*(t) = \arg \left\{ \sum_i P^i \mu_i \right\}$
is optimal solution
STOP
YES
NO
Go To 10

SUBPROBLEM

The inputs to the flow chart are,

System matrices F, G

Loss functions, $f(x,u) = f_0(x,u) + f(u)$

where $f_0(x,u) = f_0'x + g_0'u$

Parameter, ρ.

Final time T [assuming initial time 0]

Final state S $\left[\text{or } S_T - S_0^F\right]$

Polyhedral matrix, A

Right-hand side vector, b

Unit vectors, $U_0, e_1, e_2, \ldots, e_{n+1}$

$e_0 = 0$

the vector $\underline{1}$.

To retain the continuous-time aspects of the control problems, this algorithm requires explicit knowledge or availability of the matrix e^{Ft}, and its time derivatives.

It is well known [4] that the components of this $(n \times n)$ matrix can be expressed as polynomials of an order less than or equal to n with an exponential multiplying factor. The knowledge of e^{Ft} is required for determining the functions $\gamma(t)$, used in the parametric programming subproblem, and for determining the vectors P, given a control function $u(t)$, over an interval.

We can express the fundamental matrix as

$$e^{Ft} = \sum_{k=0}^{m-1} \alpha_k(t)\, F^k ,$$

where m is the degree of the minimal polynomial of F. Note that $m \leq n$, if F is $(n \times n)$. The algebraic equations determining α_k are

$$M\,\alpha = \phi(t)\ ,$$

where α is an m-vector and ϕ is a vector with elements of the form $t^{\ell} e^{s_k t}$, $k = 1, \ldots, \sigma$; $\ell = 0, \ldots, m_k - 1$, where the s_k's are the eigenvalues of F and m_k is the multiplicity of the eigenvalue, s_k. M can be shown to be nonsingular. Thus α_k is composed of linear combinations of the elements of $\phi(t)$, which are themselves members of the class of solutions to homogeneous, constant coefficient, linear differential equations. Thus any vector

$$\gamma(t) = \pi' e^{F(T-t)}\ G\ ,$$

where π is a real vector and G is a real matrix, has components which are members of the class of solutions to the homogeneous, constant coefficient, linear differential equations.

As will be proven in the next chapter, the functions generated for the linear loss case, minimum fuel problem, and minimum time problem are piecewise constant functions. Also, for the quadratic loss in control problem, the generated function $u(t)$ is shown to be expressed as a linear combination of the α_k's for a finite interval of t. Thus the components of the vector integral

$$\int_{t_1}^{t_2} e^{F(T-t)}\ Gu(t)\ dt$$

can be represented by a sum of integrals of the form,

$$\int_{t_1}^{t_2} a_s t^{\ell} e^{st}\ dt\ ,$$

which when integrated, is equivalent to

$$a_s \frac{t^{\ell} e^{st}}{s} \Bigg|_{t_1}^{t_2} - \int_{t_1}^{t_2} a_s \frac{\ell}{s}\, t^{\ell-1}\ e^{st}\ dt\ . \qquad (5.56)$$

It should be noted that all components of e^{Ft} will be real, even for complex eigenvalues, since cancellations of the complex part occur.

Since the values of e^{Ft} and its first $n - 1$ derivatives are needed at only a finite number of points (undetermined), these values may be provided by an analog computer. This suggests (but not restrictively) the use of hybrid computers for the algorithm. The analog computer could be used to supply values of $\gamma(t)$ at specified instants and to compute the vectors, $P[u(t)]$, while a digital computer could be used to solve the linear and quadratic programming problems in the master and subproblems.

To show that the algorithm is computationally feasible, we will show that each step or block in the flow chart is solvable by a finite number of iterations. Although convergence of generalized programs may be an infinite process, a suboptimal solution as close as desired to the optimal solution is achievable in a finite number of iterations of the subproblem. However, when the reachable set is a polyhedron, the generalized program converges in a finite number of iterations of the subproblem. The number of these iterations is less than or equal to the number of extreme points of the polyhedron.

To demonstrate the finiteness of the executions at each stage in the algorithm, we will show finiteness for each block of the flow chart for the basic algorithm (note that the block numbers designated coincide with those on the flow chart.) It is also noted where an analog computer may be substituted when hybrid computations are desired.

Block 1: If the matrix e^{Ft} is not available as an input, its determination may be obtained using an eigenvalue analysis of the matrix F; routines of this nature are available. After the eigenvalue analysis, a set of linear equations must be solved to find $e^{F(T-t)}$ in terms of a finite sum of multiples of F. The analog computer may be utilized for determination of e^{Ft}.

Block 2: The extreme points of a simplex, in the reachable set surrounding the desired point S, is a computation involving addition and subtraction of the available vectors, some of which are unit vectors.

Block 5 : The solution to the phase I master problem at any k for any right-hand side is a linear program that has an initial basic solution immediately available, i.e.,

$$\mu_i = 0, \ \forall i$$

$$y_i^+ = 0, \ y_i^- = -S_i \ , \qquad \text{for} \ S_i < 0$$

$$y_i^+ = S_i, \ y_i^- = 0 \ , \qquad \text{for} \ S_i \geq 0 \ .$$

The number of rows in this linear program is $n + 1$ for any stage k, even though the number of columns is variable but always finite.

Block 6 : Additional columns are added to the master program of phase I until the value of w is less than a required strictly positive number. Since the minimum value of w is zero and since w decreases monotonically and strictly decreases on a subsequence of iterations, for some specified positive number, the value of w will be smaller than this number after a finite number of iterations. This is a basic result of generalized programming problems and shows only a finite number of columns are used for the master program.

Block 7 : The calculation of the dual variable of the linear program in the master problem is a result of the solution procedure for the linear program and requires little or no additional computation.

Block 8 : A vector addition provides the column vectors to be used in the master program after the phase I procedure is completed.

Blocks 9 and 13 : The determination of the vector $\overline{P}$ is achieved by intregation. However, due to the structure of the integrand, specialized (finite and exact) integration schemes are possible. The integral is broken into a finite sum of definite integrals (corresponding to a finite number of switching points) whose end points are calculated by the analog computer or by formula substitution as suggested by Eq. (5.56).

Block 11 : The master program for the optimization phase is a linear program with a fixed number of rows (at most $n + 3$) and a variable number of columns. Although the number of columns may be infinite, for any practical problem and within limits of the computer's accuracy, no more than a finite number of columns are generated before achieving a solution, indistinguishable (within computer accuracy) from the optimal solution. The simplex method should be used with the starting solution to each interation being the final solution of the previous iteration. The first vector to be added to the basis is the vector generated from the subproblem (if optimality has not already been achieved).

Block 12 : The solution of the parametric programming problem is discussed in Chapter IV. The solution has a finite number of executions; cycling is avoided due to the lexicographic ordering rules and normal degeneracy perturbation techniques available for linear programming codes and complementary pivot theory methods.

All other steps are either logical programming steps or simple calculations. Thus since each step requires a finite number of executions, each iteration of the master problem and its corresponding subproblem (of phases I and II) require a finite number of executions.

The solution of phase I is finite since only a finite number of columns must be generated for the $n + 1$ generalized programs used for the solution to each step of phase I. The algorithm may be terminated at any stage in phase II yielding a feasible control with a bound on how much its objective value can differ from the optimal objective value.

Chapter VI

RELATION OF GENERALIZED PROGRAMMING TO CONTROL THEORY

In this chapter, the relationship between the necessary conditions of the generalized programming formulation and Pontryagin's necessary conditions for the optimal control problem is discussed. The characteristics of optimal controls for the various classes of control problems are also discussed.

A. Relation to Pontryagin's Necessary Conditions

The relationship between the generalized programming optimality conditions and Pontryagin's necessary conditions is used to show how a solution to the generalized program can be an optimal solution to the control problem. The following class of problems (discussed in the previous chapter) are considered:

$$x \in E^n , \qquad u \in E^m$$

$$\dot{x} = Fx + Gu$$

$$\min J = \int_0^T \left\{ f_0'x + g_0'u + f(u) \right\} dt ,$$

where $$f(u) = \begin{cases} 0 \\ \sum |u_i| , \\ u'Qu \end{cases}$$

and $$x(0) = 0$$

$$x(T) = S$$

$$u(t) \in U = \{u | Au \geq b\}$$

$$\forall t .$$

If we let $\dot{x}_0 = f_0'x + g_0'u + f(u)$, then

$$\dot{\bar{x}} = \begin{bmatrix} \dot{x}_0 \\ \dot{x} \end{bmatrix} = \overline{F}\overline{x} + \overline{G}u + f(u)\, U_0 .$$

By using the previously presented notation, the Hamiltonian can be defined as

$$H = \bar{\psi}'\bar{F}\bar{x} + \bar{\psi}'\bar{G}u + \bar{\psi}' f(u)\ U_0 \tag{6.1}$$

$$\dot{\bar{x}} = H_{\bar{\psi}} , \quad \text{and} \tag{6.2}$$

$$\dot{\bar{\psi}} = -H_{\bar{x}} , \tag{6.3}$$

$$\text{where} \quad \bar{\psi}(t) = \begin{bmatrix} \psi_0(t) \\ \\ \psi(t) \end{bmatrix} \quad \text{and} \quad \psi(t) = \begin{bmatrix} \psi_1(t) \\ \vdots \\ \psi_n(t) \end{bmatrix} .$$

Pontryagin's maximum principle states that (6.2) and (6.3) must hold while the optimum $u^*(t)$ satisfies

$$H(\bar{\psi}, \bar{x}, u^*) = \sup_{u(t) \in U} H(\bar{\psi}, \bar{x}, u) , \tag{6.4}$$

for a given ψ, x, or

$$H(\bar{\psi}, \bar{x}, u^*) \geq H(\bar{\psi}, \bar{x}, u) , \quad \text{all} \quad u(t) \in U . \tag{6.5}$$

Let the optimal dual solution to the generalized program,

$$\max_{P \in C} \lambda$$

$$\text{subject to} \quad U_0\lambda + \bar{P}\mu = \bar{S}$$

$$\mu = 1$$

$$\bar{P} \in \bar{C} ,$$

be

$$\bar{\pi}^{*\prime} = (\pi_0^*, \pi_1^*, \ldots, \pi_{n+1}^*) ;$$

and let

$$\bar{\psi}(t) = -\tilde{\pi}^{*\prime}\ e^{\bar{F}(T-t)} , \tag{6.6}$$

where $\tilde{\pi}^{*\prime} = (\pi_0^*, \ldots, \pi_n^*)$. Thus it is obvious that $\overline{\psi}(t)$, as defined, provides a solution to (6.3). To show that this solution is non-trivial, it is sufficient to show that $\psi_0(t) \neq 0$. From the results obtained in Chapter V, the first column of $e^{\overline{F}(T-t)}$ is the unit vector U_0. Thus,

$$\psi_0(t) = -\pi_0^* , \qquad \text{for all } t . \tag{6.7}$$

We also know, from generalized programming, that

$$\tilde{\pi}^* U_0 = 1 ,$$

therefore, $\pi_0^* = 1$ and $\psi_0(t) = -1$, for all t.

Since $\overline{\pi}^*$ is the dual solution to the generalized program, it satisfies

$$\min_{\overline{P} \in \overline{C}} \overline{\pi}^{*\prime} \begin{bmatrix} \overline{P} \\ 1 \end{bmatrix} \leq \overline{\pi}^{*\prime} \begin{bmatrix} \overline{P} \\ 1 \end{bmatrix} , \qquad \overline{P} \in \overline{C} ,$$

or, equivalently,

$$\min_{u(t) \in U} \overline{\pi}^{*\prime} \begin{bmatrix} \overline{P}(u) \\ 1 \end{bmatrix} \leq \overline{\pi}^{*\prime} \begin{bmatrix} \overline{P}(u) \\ 1 \end{bmatrix} \qquad \forall u(t) \in U . \tag{6.8}$$

The above inequality is equivalent to the subproblem of the generalized program, when $\overline{\pi}^*$ is the current dual variable. Equation (6.8) may be restated as

$$\pi_0^* f(u^*) + \tilde{\pi}^{*\prime} e^{\overline{F}(T-t)} \overline{G}u^*(t)$$

$$\leq \pi_0^* f(u) + \tilde{\pi}^{*\prime} e^{\overline{F}(T-t)} \overline{G}u(t)$$

for all $u(t) \in U$, and

$t \in [0,T]$,

or

$$-\pi^{*} f(u^{*}) - \tilde{\pi}^{*'} e^{\bar{F}(T-t)} \bar{G}u^{*}(t)$$

$$\geq -\pi_0^{*} f(u) - \tilde{\pi}^{*'} e^{\bar{F}(T-t)} \bar{G}u(t) .$$

By expressing the above in terms of $\bar{\psi}(t)$ and by using Eqs. (6.6) and (6.7), the inequality is equivalent to (6.5) or Pontryagin's maximum principle. Thus the generalized programming necessary conditions are equivalent to Pontryagin's necessary conditions for the same problem.

To complete the analogy between Pontryagin's necessary conditions and the generalized programming optimality conditions for control problems, we include here a discussion of the transversality conditions for sets $\mathcal{S}$ which are convex smooth manifolds.

Notice that the vector $\tilde{\pi}^{*}$ is equivalent to the vector $\bar{\psi}(T)$. Also an optimality condition for free (initial) final point problems [as shown in Eq. (5.47)] is

$$\bar{\pi}^{*'} \begin{bmatrix} 0 \\ S \\ 0 \\ 1 \end{bmatrix} \geq 0 , \qquad \text{for all} \quad S \in \mathcal{S} .$$

We also note that some $S^{*} \in \mathcal{S}$ has the property

$$\pi^{*'} S^{*} = -\pi^{*}_{n+2} ,$$

since there must be some vector S^{*} in the basis of the expanded master linear program.

These conditions represent a halfspace with the hyperplane defining it as being represented by the vector $\bar{\pi}^{*}$. This hyperplane is a supporting hyperplane to $\mathcal{S}$ at some S^{*}, since $\pi^{*'} S^{*} = -\pi^{*}_{n+2}$; $\mathcal{S}$ lies completely in one halfspace of the hyperplane. The hyperplane is also a tangent plane to the manifold $\mathcal{S}$ (when $\mathcal{S}$ is a manifold). These conditions represent the fact that $\bar{\psi}(t) = -\tilde{\pi}^*$ is orthogonal to the tangent hyperplane of $\mathcal{S}$ at S^*. This is precisely the transversality condition described in Chapter II.

To show that a solution to the generalized programming problem is also a solution to the continuous-time optimal control problem (for fixed end points), it is assumed that we have a finite set of vectors $\bar{P}^i$, so that

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P}^i \\ 1 \end{bmatrix} = 0 \qquad \text{and}$$

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} \geq 0, \qquad \text{all } \bar{P} \in \bar{C} .$$

Also

$$\sum \bar{P}^i \mu_i = \bar{S}$$

$$\sum \mu_i = 1$$

$$\mu_i \geq 0 \tag{6.9}$$

has a solution, μ_i^*.

Theorem 6.1. The solution

$$u^*(t) = \sum u^i(t)\, \mu_i^*$$

$$t \in [0,T]$$

is an optimal solution for the control problem.

Proof of Theorem 6.1.

We know

$$\bar{\psi}(t) = -\tilde{\pi}^{*\prime} e^{\bar{F}(T-t)}$$

satisfies (6.3) and $\bar{\psi}_0(t) = -1$ satisfies $\bar{\psi}_0(t) \leq 0$, $t \in [0,T]$.

$$f(u^*) + \tilde{\pi}^{*\prime} e^{\bar{F}(T-t)} \bar{G}u^*(t) = f\left[\sum \mu_i^* u^i(t)\right]$$

$$+ \tilde{\pi}^{*\prime} e^{\bar{F}(T-t)} \bar{G}\left[\sum \mu_i^* u^i(t)\right] ,$$

and, from convexity,

$$f(u^*) \leq f\left[\sum \mu_i^* u^i(t)\right] \leq \sum \mu_i^* f(u^i) .$$

Since $\delta^k = 0$, from generalized programming necessary conditions,

$$0 = f(u^*) + \tilde{\pi}^{*\prime} e^{\bar{F}(T-t)} \bar{G}u^*(t) + \pi^*_{n+1}$$

$$\leq \sum \mu_i^* f(u^i) + \sum \mu_i \tilde{\pi}^{*\prime} e^{\bar{F}(T-t)} \bar{G}u^i(t) + \sum \mu_i^* \pi^*_{n+1}$$

$$= \sum \mu_i^* \bar{\pi}^{*\prime} \begin{bmatrix} \bar{P}^i \\ 1 \end{bmatrix} = 0 .$$

Since π^*_{n+1} is constant and since

$$\bar{\pi}^* \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} \geq 0 , \qquad \forall \bar{P} \in \bar{C} ,$$

$$f(u^*) + \tilde{\pi}^{*'} e^{F(T-t)} \bar{G}u^*(t) \leq f(u) + \tilde{\pi}^{*'} e^{\bar{F}(T-t)} \bar{G}u(t),$$

$$\forall u(t) \in U ,$$

or

$$- f(u^*) - \tilde{\pi}^{*'} e^{\bar{F}(T-t)} \bar{G}u^*(t) \geq -f(u) - \tilde{\pi}^{*} e^{\bar{F}(T-t)} \bar{G}u(t),$$

$$u(t) \in U ,$$

$$t \in [0,T] ,$$

which is equivalent to the maximum principle. Thus Pontryagin's necessary conditions for optimality are satisfied by the solution to the generalized programming problem. Since the set of equations (6.9) has a solution, the control $u^*(t)$ is a feasible control. It remains to be shown that

$$J(u^*) \leq J(u) , \qquad \forall u(t) \in U$$

which was shown to be a result of $\delta^k = 0$, in Theorem 5.6. Thus the control $u^*(t)$ is an optimal control for the continuous-time control problem.

Q.E.D.

We will now show that given $\bar{\pi}^*$, an optimal dual solution, a finite set of vectors $\bar{P}^i$ can be found to provide a feasible solution to the set of equations (6.9). This is done for three cases, the quadratic loss in control with positive definite matrix Q, the linear cases satisfying Pontryagin's general position condition, and finally, the linear cases not satisfying Pontryagin's general position condition.

<u>Theorem 6.2.</u> If the quadratic loss control problem with positive definite Q has a feasible solution, then $u^*(t)$, which provides a solution to

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} = 0 \, ,$$

is an optimal control for the control problem.

Proof of Theorem 6.2.

Let $u^*(t)$ be a control satisfying

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} = 0 \, ,$$

then $u^*(t)$ solves

$$\min \; u(t)' \, Qu(t) + \gamma^*(t)' \, u(t)$$

$$u(t) \in U \, , \qquad t \in [0,T] \, ,$$

where $\gamma^*(t)$ is generated by $\bar{\pi}^*$. The quadratic program with Q positive definite has a unique solution at each t. Thus there is no other $u(t)$ satisfying

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P}(u) \\ 1 \end{bmatrix} = 0 \, .$$

By the feasibility assumption and by U being a compact set, an optimal control exists and must satisfy the necessary conditions. Since $u^*(t)$ is the only control satisfying the necessary conditions for optimality, it must be the optimal control.

Q.E.D.

Theorem 6.3. For feasible linear control problems (including minimum fuel and minimum time) where $\bar{F}$, $\bar{G}$, and U satisfy the general position condition of Pontryagin, the control $u^*(t)$ satisfying

$$\bar{\pi}^{*'} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} = 0$$

is an optimal control.

Proof of Theorem 6.3.

Since a feasible control exists and since U is compact, an optimal control $u^{**}(t)$ must exist and satisfy the necessary conditions for optimality. From the results of Pontryagin, when the general position condition holds,

$$\min_{u(t) \in U} \bar{\pi}^{*'} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} \tag{6.10}$$

has a unique solution [of u(t)], except on a set of measure zero, which we call $u^{*}(t)$. Thus $u^{*}(t)$ is the only function satisfying the necessary conditions, and

$$u^{*}(t) \equiv u^{**}(t)$$

Q.E.D.

When the general position condition is not satisfied by $\bar{F}$, $\bar{G}$, and U, the solution to (6.10) is not necessarily unique over a set of positive measure. However, since (6.10) must be satisfied (because it is a necessary condition), only its solutions need be investigated to produce the optimal control. This is true, since the problem is feasible, U is compact, and an optimal control exists. By the theory of generalized programming, any solution to (6.10) which is feasible for (6.9) is an optimal solution, as shown by Theorem 6.1.

Proposition 6.1. There are a finite number of distinct solutions to (6.10).

Proof of Proposition 6.1.

It has been shown in Chapter IV that an upper bound exists on the number of possible switching points of any solution to the parametric programming problem (6.10), for any value of $\bar{\pi}^*$ for a finite interval of t. These points are fixed (given $\bar{\pi}^*$), and any solution to (6.10) remains constant between any neighboring pair of switching points. There are a finite number of possible solutions for (6.10), between such switching points (due to a finite number of bases). Thus there are a finite number of distinct solutions to (6.10).

Q.E.D.

Proposition 6.2. The set of points $\bar{P}$, generated by (6.10), are extreme points of a convex (bounded) polyhedron of all $\bar{P}$ satisfying

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} = 0 . \qquad (6.11)$$

Proof of Proposition 6.2.

By the minimization procedure of (6.10), extreme points are generated. Also by the homogeneity of (6.11), any convex combination of the finite number of extreme points satisfy (6.10). Thus the points generated by (6.10) are extreme points of the convex polyhedral set containing the solutions to (6.11).

Q.E.D.

Theorem 6.4. For linear loss problems not satisfying the general position condition, a finite set of vectors $\bar{P}^i$ can be found to provide an optimal solution to the generalized programming problem.

Proof of Theorem 6.4.

The optimal solution satisfies

$$\bar{\pi}^{*\prime} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} = 0 ,$$

hence we have shown that a finite number of vectors $\overline{P}$ form the extreme points of all vectors satisfying the Eq. (6.11). Since an optimal solution exists, it must satisfy (6.10), and therefore, it must be a convex combination of all extreme points to the set of vectors satisfying (6.11). It can be shown that an optimal control vector P^* must be a combination of at most $n + 1$ extreme vectors $\overline{P}^i$ satisfying (6.10). Thus, the finite set of extreme vectors satisfying (6.10) can be generated and must include, in its convex hull, a feasible solution to (6.9). Therefore, the optimal solution is determined by a finite set of vectors $\overline{P}^i$.

Q. E. D.

Let μ_i^* be the solution to

$$U_0\lambda + \mu_1^*\overline{P}^1 + \dots + \mu_k^*\overline{P}^k = \overline{S}$$

$$\sum_{i=1}^{k} \mu_i^* = 1 \ , \qquad (6.12)$$

where the $\overline{P}^i$ are the extreme vectors of (6.10) or (6.11). Such a solution exists from Theorem 6.4.

Theorem 6.5. $u^*(t) = \sum_{i=1}^{k} \mu_i^* u^i(t)$,

where $u^i(t)$ generates P^i, is an optimal solution to the control problem.

Proof of Theorem 6.5.

$u^*(t)$ provides a feasible solution by virtue of (6.12) and satisfies the necessary conditions by construction; by Theorem 6.1, the cost function is minimal over all feasible controls.

Q. E. D.

B. Characteristics of Linear Loss Optimal Controls

Without assuming a general position condition, we will show that an optimal solution to the control variable for line r loss functionals, including minimal fuel and minimum time problems is a piecewise constant function with a finite number of points o discontinuity for any finite interval of time. Thus, since we are considering finite horizon problems $(T < \infty)$, an optimal control is a piecewise constant vector function with a finite number of switching points.

The solution to the parametric linear program is observed as being a vector control function that is piecewise constant and has a finite number of switching points. Thus any vector P^i, generated by the subproblem, has the same property for its generating control function. Since generalized programming problems are linear programs in the master problem, and since the number of rows in the linear program is less than or equal to $(n + 3)$, the number of columns P^i in any solution is at most $(n + 1)$ for phases I and II of the algorithm.

Proposition 6.3. The columns P^i, generated for the optimization phase of the algorithm by the phase I procedure, are generated by control vectors that are piecewise constant with a finite number of discontinuities.

Proof of Proposition 6.3.

In general, the $n + 1$ columns P^i, for $i = 0, 1, \ldots, n$ generated for an initial feasible solution to the control problem, are generated from the phase I algorithm for $n + 1$ right-hand sides. Thus each P^i is generated by a set of at most $n + 1$ vectors P, and each in turn is generated by a piecewise constant control function with a finite number of switchings. If the maximum number of switchings for any control function generated by the subproblem is M $(M < \infty)$, each P^i has a control with at most $(n + 1)M$ switchings, since these controls are generated by summing $n + 1$ control functions with at most M switchings each.

Each new column (after feasibility is attained) is generated by a piecewise constant control with at most M switchings. Thus each column P^i, in the master problem of phase II, has at most $(n + 1)M$ switchings in control.

Q.E.D.

<u>Proposition 6.4.</u> At any stage in the iterative process of the generalized program, the current control solution is a piecewise constant function with a finite number of discontinuities and has an objective value within the bound of the optimal objective value, given in Eq. (5.53), of δ^k.

<u>Proof of Proposition 6.4.</u>

The solution to the master problem contains a nonnegative combination of at most $n + 1$ columns P^i, each generated by a control function that is piecewise constant, and each has at most $(n+1)M$ switching points (Proposition 6.3). Thus, the combination of the controls to generate the solution has at most $(n + 1)^2M$ discontinuities. The bound (δ^k) was shown in Chapter V.

Q.E.D.

<u>Theorem 6.6.</u> The optimal control generated by the generalized programming solution of the continuous-time problem for the linear loss functionals (minimum fuel and minimal time problems included as special cases) is a piecewise constant function, and it has a finite number of discontinuities.

<u>Proof of Theorem 6.6.</u>

If the generalized program terminates with a value of $\delta^k = 0$ for some stage k, then by Proposition 6.4, the theorem is true.

In any case, given the optimal dual variables to the generalized program $\bar{\pi}^*$, the optimal solution is a combination of at most $n + 1$ vectors P^i, generated as solutions to (6.10). The generating controls of these vectors have at most M switchings, and their combinations has most $(n +1)M$ switchings.

Q.E.D.

It should be noted that when $\delta^k = 0$, the current value of $\bar{\pi}^k$ is optimal, and it can be used to determine an optimal control function independent of the current solution (but not necessarily distinct).

Theorem 6.7. If the solution to the parametric linear programming problem when using $\bar{\pi}^*$ to generate $/(t)$ is unique except on a set of measure zero, the optimal control function, when non-zero, is at an extreme point of the admissible control region.

Proof of Theorem 6.7.

The solution to a linear program always occurs at an extreme point of the constraint set. When a change of variables is made to produce an equivalent problem for minimum fuel problems, a control of level zero is considered to be at an extreme point of the new constraint set. Thus, the optimal control is at an extreme point of the admissible control region (or an equivalent constraint set for minimum fuel or minimum time problems).

Q.E.D.

The previous theorem also implies that the standard minimal time problem, and certain linear loss problems, have bang-bang solutions. It also implies that the minimal fuel solution is a bang-coast-bang solution in some cases.

Pontryagin [1] has shown that his general position condition is a sufficient condition to insure that the parametric linear programming problem has a unique solution almost everywhere.

Proposition 6.5. The upper bound on the number of switchings for the linear loss functionals when the matrix $\bar{F}$ has real eigenvalues, the state of the system is n, and the matrix A of the admissible control region is $m \times p$, is

$$(n + 1)^2 (p - m) \binom{p}{m} .$$

Proof of Proposition 6.5.

There are at most $\binom{p}{m}$ bases for the parametric linear program. Each has $(p - m)$ nonbasic variables with relative cost factors $\bar{\gamma}_i(t)$ having at most $(n + 1)$ points at which it becomes value zero. Thus if a basis can be repeated, it can do so no more than $(n + 1)(p - m)$ times, after which it remains optimal. Thus each column P^i of the master program is generated by a control with at most

$$(n + 1)(p - m)\binom{p}{m}$$

switching points. Since at most $(n + 1)$ control functions are combined, the maximum number of switchings is

$$(n + 1)^2 (p - m)\binom{p}{m}.$$

Q.E.D.

C. Characteristics of Quadratic Loss Optimal Controls

Since the parametric quadratic program has a time (parameter) dependent solution for the control vector, the only characterization of the optimal control generated is in the class of time functions possible for the solution.

For the quadratic programming problem, stated as

$$w = Mz + q(t) \ , \quad w, z \geq 0, \quad w_i z_i = 0 \ ,$$

the solution has the form

$$\begin{bmatrix} \bar{w} \\ \bar{z} \end{bmatrix} = \bar{q}(t) + \bar{M} \begin{bmatrix} \underline{z} \\ \underline{w} \end{bmatrix},$$

where

$$\begin{bmatrix} \bar{w} \\ \bar{z} \end{bmatrix} = \bar{q}(t), \quad \text{and} \quad \underline{z}, \underline{w} = 0 \ .$$

$\bar{q}_i(t)$ is merely a linear combination of the original components of $q(t)$ which are again linear combinations of the components

$$t^{\ell} e^{s_k t}$$

for all s_k eigenvalues of $\bar{F}$, and ℓ less than the multiplicity of the eigenvalue s_k. Thus the solution to the parametric quadratic program is composed of linear combinations of the same elements.

As in the linear case, at most $(n + 1)$ columns of the generalized programming master problem are used at any stage; therefore, the current solution is of the same form, i.e., linear combinations of the elements

$$t^{\ell} e^{s_k t} .$$

Whenever a basis switch is made, the linear multipliers change in the linear combination, but the solution has the same characteristic form.

Proposition 6.6. At any stage in the quadratic control problem, including the optimal solution, the form of each component of the control function is a linear combination of $n + 1$ terms of the form

$$t^{\ell} e^{s_k t} ,$$

with only the constant terms changing at each of the finite number of basis switches.

Proof of Proposition 6.6.

Each column generated by the subproblem has a control function of the required form with the finite number of basis switches, since the control function is generated from the parametric quadratic program. Since at most $(n + 1)$ columns are combined for each solution, the current control solution has the same form. At the optimum, the quadratic program has a unique solution for positive definite Q, and only one column is generated with the generating control, which is optimal, having the required form. For positive semidefinite Q, the solution to the quadratic program is not necessarily unique, and a combination of at most $(n + 1)$ control functions may be required as in the linear case.

Q.E.D.

This algorithm provides an open loop solution to the optimal control problem. Also, it should be noted that no assertion is made regarding the uniqueness of the solution in the form of the optimal control function.

Since Chapter VII provides an illustrative example with computational experience in the linear case, we will now present an example showing the form of the solution for the case of quadratic loss in control [14].

Consider the system

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = u .$$

Thus,

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$x_1(0) = 0, \qquad x_2(0) = 0$$

$$x_1(T) = s_1 , \qquad \text{and} \quad x_2(T) = s_2 .$$

$$\min_{u(\cdot)} \quad J = \int_0^T \frac{1}{2} u^2 \, dt$$

$$|u| \leq 1 ,$$

where $Q = \frac{1}{2} I$ and is positive definite. If the optimum dual variable to the generalized programming formulation of this problem is

$$\bar{\pi}^{*'} = (-1, \pi_1^*, \pi_2^*, \pi_3^*) ,$$

the optimal solution can be obtained as follows:

find the control $u^*(t)$ so that

$$\min_{|u(t)| \leq 1} \bar{\pi}^{*'} \begin{bmatrix} \bar{p} \\ 1 \end{bmatrix} \quad \text{is achieved by } u^*(t) .$$

Therefore, $u^*(t)$ must solve

$$\min_{u(\cdot)} \left\{ \frac{1}{2} u^2 + (\pi_1^*, \pi_2^*) e^{F(T-t)} Gu(t) \right\} + \pi_3^*$$

$$|u(t)| \leq 1, \quad t \in [0,T] \text{ , or}$$

$$\min_{u(\cdot)} \left\{ \frac{1}{2} u^2 + \left[\psi_1^*(t), \psi_2^*(t)\right] Gu(t) \right\}$$

$$|u(t)| \leq 1 ,$$

where $\psi_1^*(t)$, and $\psi_2^*(t)$ are the optimal adjoint variables for all t. Hence the minimization is

$$\min_{u(\cdot)} \left\{ \frac{1}{2} u^2 + \psi_2^*(t)' u \right\}$$

$$|u(t)| \leq 1..$$

For this problem, the solution is easily seen as

$$u^*(t) = -\text{sat } \psi_2^*(t) = \begin{cases} -\psi_2^*(t), & |\psi_2^*(t)| \leq 1 \\ -\text{sgn } \psi_2^*(t), & |\psi_2^*(t)| > 1 . \end{cases}$$

Note that $\psi_2^*(t)$ has the form

$$(\pi_1^*, \pi_2^*)\begin{bmatrix} T-t \\ 1 \end{bmatrix} = \pi_1^*(T-t) + \pi_2^* , \qquad \text{i.e.,}$$

$$\psi_2^*(t) = \alpha t + \beta .$$

Thus the optimal solution has the form $u(t) = \alpha t + \beta$, for any interval where α and β are allowed to change at certain switching points.

Free Final State Problem. In conclusion we will consider the problem where the initial state is zero and the final state (at fixed time, T) is completely free. Since the first through n^{th} rows of the generalized programming master problem have free right-hand sides, the slack variables for these rows are always permitted to be non-zero. Thus, the optimal $\bar{\pi}^*$ must be

$$\bar{\pi}^{*\prime} = (1, 0, \ldots, \pi_{n+1}^*) .$$

The optimal control is then determined by the subproblem,

$$\min_{u(\cdot)} f[x(u), u]$$

$$\text{subject to} \quad u(t) \in U \qquad t \in [0,T] .$$

Thus the problem is solved by one iteration of a parametric programming problem.

Chapter VII

EXAMPLES AND COMPUTATIONS

In this chapter, examples are used to show the execution and sample results of the algorithm. The convergence properties will be demonstrated as well as the basic features of the algorithm.

The problem we intend to solve is

$$\min_{u(\cdot)} \int_0^3 |u(t)| \, dt$$

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = u$$

$$|u(t)| \leq 1$$

$$x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad x(3) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$e^{Ft} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}, \qquad \text{and}$$

$$e^{F(T-t)} G = \begin{bmatrix} 1 & T-t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} T-t \\ 1 \end{bmatrix}.$$

Therefore,

$$P = \int_0^3 \begin{bmatrix} T-t \\ 1 \end{bmatrix} u(t) \, dt \;;$$

and

$$U = \left\{ u \,\middle|\, |u| \leq 1 \right\} .$$

By the definitions given in Chapter V,

$$S_0^F = \left\{ e^{FT} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} ,$$

$$S_T = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

$$\mathcal{S} = \{S\} = \left\{ S_T - S_0^F \right\} = \left\{ \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\} .$$

We initiate by looking for a feasible solution to

$$\min_{\mu, y} \quad w = \sum_{i=1}^{2} y_i^+ + \sum_{i=1}^{2} y_i^-$$

$$\text{subject to} \qquad P\mu + Iy^+ - Iy^- = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\mu \qquad\qquad = 1 , \qquad \text{where}$$

$$P \in C = \left\{ P \,\middle|\, P = \int_0^3 \begin{bmatrix} T-t \\ 1 \end{bmatrix} u(t)\, dt, \right.$$

$$\left. |u(t)| \leq 1 \right\} .$$

Let $P^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, generated by the admissible control $u(t) = 0$. In this example, we seek to solve phase I by keeping the desired point as the right-hand side and to terminate when $w = 0$.

The first master problem is represented by the tableau

$$\min_{\mu, y} w = \sum_{i=1}^{2} y_i^+ + \sum_{i=1}^{2} y_i^-$$

μ_0	y_1^+	y_2^+	y_1^-	y_2^-		
0	1	0	-1	0	=	-1
0	0	1	0	-1	=	0
1	0	0	0	0	=	1

$$\mu_0,\ y_i^+,\ y_i^- \geq 0 \ .$$

The first two rows correspond to states in the dynamic system, and the final row represents the possibilities of convex combinations of the columns P^i, generated by the subproblem. The optimal basis for this linear program is

$$B_0 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

with

$$B_0^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} .$$

The solution is

$$\mu_0 = 1,\ y_1^- = 1,\ y_1^+ = y_2^+ = y_2^- = 0$$

with

$$w^0 = 1 \ .$$

The dual variable is

$$\hat{\pi}^{0'} = \gamma_B B^{-1} = (0,\ 1,\ 1)\ B^{-1} = (-1,\ 1,\ 0)\ ,$$

$$[\hat{\pi}' = (\pi_1,\ \pi_2,\ \pi_3)]\ .$$

The subproblem for the first interation is

$$\min_{P \in C} - \hat{\pi}' \begin{bmatrix} P \\ 1 \end{bmatrix} .$$

(The minus sign results from the manner in which π is generated from γ_B.) The subproblem is then expanded to

$$\min_{u(t) \in U} \int_0^3 (-\pi_1,\ -\pi_2) \begin{bmatrix} 3-t \\ 1 \end{bmatrix} u(t)\ dt - \pi_3$$

or

$$\min_{u(\cdot)} [-\pi_1(3 - t) - \pi_2]\ u(t),$$

$$|u(t)| \leq 1\ , \qquad t \in [0,3]\ ,$$

$$\text{for} \quad \pi_1 = -1,\ \pi_2 = 1\ .$$

The minimum is achieved by the function

$$u^1(t) = -1 \qquad t \in [0,2]$$

$$u^1(t) = 1 \qquad t \in (2,3]$$

which generates a vector P^1,

$$P^1 = \int_0^3 \begin{bmatrix} 3-t \\ 1 \end{bmatrix} u(t)\,dt = \begin{bmatrix} -3.5 \\ -1.0 \end{bmatrix} .$$

The new master problem tableau has an additional column corresponding to P^1, i.e.,

$$\min_{\mu, y} w = \sum_1^2 y_i^+ + \sum_1^2 y_i^-$$

μ_0	μ_1	y_1^+	y_2^+	y_1^-	y_2^-		
0	-3.5	1	0	-1	0	=	-1
0	-1.0	0	1	0	-1	=	0
1	1	0	0	0	0	=	1

$$\mu_i,\ y_i^+,\ y_i^- \geq 0 .$$

The optimal basis for this linear program is

$$B_1 = \begin{bmatrix} -3.5 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} ;$$

and the solution is

$$\mu_0 = 5/7, \quad \mu_1 = 2/7, \quad y_2^+ = 2/7, \quad \text{all others equal zero,}$$

$$w^1 = 2/7 .$$

Using the dual variable from the above basis, the new column generated by both the subproblem and the generating control is

$$p^2 = \begin{bmatrix} 4.5 \\ 3 \end{bmatrix}, \quad u^2(t) = 1, \quad t \in [0,3] .$$

The new tableau is

$$\min_{\mu, y} \quad w = \sum_i^2 y_i^+ + \sum_i^2 y_i^-$$

μ_0	μ_1	μ_2	y_1^+	y_2^+	y_1^-	y_2^-		
0	-3.5	4.5	1	0	-1	0	=	-1
0	-1	3	0	1	0	-1	=	0
1	1	1	0	0	0	0	=	1

$$\mu_i, \; y_i^+, \; y_i^- \geq 0 .$$

The optimal basis for this linear program is

$$B_2 = \begin{bmatrix} 0 & -3.5 & 4.5 \\ 0 & -1 & 3 \\ 1 & 1 & 1 \end{bmatrix},$$

and the solution is

$$\mu_0 = 1/3, \quad \mu_1 = 1/2, \quad \mu_2 = 1/6$$

with

$$w^2 = 0 .$$

Thus, a feasible control has been found for the control problem.

To demonstrate the above fact, let

$$u^f(t) = \mu_0 u^0(t) + \mu_1 u^1(t) + \mu_2 u^2(t) \ , \qquad \text{or}$$

$$u^f(t) = \begin{cases} -\frac{1}{3} & t = [0,2] \\ \frac{2}{3} & t = (2,3] \end{cases} \ ;$$

and let

$$P^f = P[u^f(t)] = \int_0^3 \begin{bmatrix} 3-t \\ 1 \end{bmatrix} u^f(t)\, dt = \begin{bmatrix} -1 \\ 0 \end{bmatrix} .$$

Thus $u^f(t)$ is a feasible control, i.e., an admissible control bringing the system from its initial state to the desired final state. Note that

$$J[u^f(t)] = \int_0^3 |u^f(t)|\, dt = 4/3 \ .$$

Since a feasible solution is available consisting of a positive combination of $(n + 1)$ vectors P, the optimization phase may be initiated.

Let

$$\overline{P}^i = \begin{bmatrix} J[u^i(t)] \\ P^i \end{bmatrix} ; \qquad \text{therefore}$$

$$p_0^0 = 0, \quad p_0^1 = 3, \quad \text{and} \quad p_0^2 = 3 \ .$$

The initial master problem is $(k = 2)$,

$$\max_{\mu} \lambda$$

$$U_0\lambda + \sum_{i=0}^{k} \bar{P}^i \mu_i = \bar{S}$$

$$\mu_i = 1$$

$$\mu_i \geq 0 \ ,$$

or in tableau form,

$$\max \lambda$$

λ	μ_0	μ_1	μ_2		
1	0	3	3	=	0
0	0	-3.5	4.5	=	-1
0	0	-1	3	=	0
0	1	1	1	=	1

$$\mu_i \geq 0 \ .$$

The solution to this linear program is

$$\mu_0 = 1/3, \quad \mu_1 = 1/2, \quad \mu_2 = 1/6 \ ,$$

with

$$\lambda^k = 2 \ .$$

If we define $\bar{u}^k$ as the control

$$\bar{u}^k = \sum_{i=0}^{k} \mu_i^k u^i(t) \ ,$$

where μ_i^k is the solution to the k^{th} master program, the inequalities in Theorem 5.6 become

$$J(u^*) \leq J(\bar{u}^k) \leq -\lambda^k = J^k .$$

For $k = 2$ the solution is

$$J(\bar{u}^k) = 4/3 , \qquad -\lambda^k = 2 .$$

The dual variable $\bar{\pi}^2$ is

$$\bar{\pi}^{2'} = (1, 2, -4, 0) ,$$

and the subproblem is

$$\text{find } \delta^k, \ u^{k+1}, \ p^{k+1}, \qquad \text{so that}$$

$$\delta^k = \min_{P \in C} \bar{\pi}^{k'} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} \qquad \text{or}$$

$$\min_{u(t) \in U} \int_0^3 \pi_0^k \, |u(t)| + \left(\pi_1^k, \pi_2^k\right) \begin{bmatrix} 3-t \\ 1 \end{bmatrix} u(t) \, dt + \pi_3^k .$$

This is equivalent to finding the solution to

$$\min_{u(\cdot)} |u(t)| + \left[\pi_1^k(3-t) + \pi_2^k\right] u(t)$$

$$|u(t)| \leq 1$$

$$t \in [0,3]$$

For $k = 2$, the solution is

$$u^3(t) = \begin{cases} -1 & t \in [0,\ 1/2] \\ 0 & t \in (1/2,\ 3/2] \\ 1 & t \in (3/2,\ 3] \end{cases}$$

This solution produces a new vector

$$\overline{P}^3 = \begin{bmatrix} 2.0 \\ -0.25 \\ 1.0 \end{bmatrix}$$

with

$$\delta^k = \overline{\pi}^{k\,\prime} \begin{bmatrix} \overline{P} \\ 1 \end{bmatrix} = -2.5 \ .$$

Thus, by using the inequalities in Theorem 5.6, we arrive at

$$J(\overline{u}^k) + \delta^k \leq J(u^*) \leq J(\overline{u}^k) \leq J^k \ ;$$

for $k = 2$, the solution is

$$-1.17 \leq J(u^*) \leq 1.33 \ .$$

After each column is generated, the master problem is augmented (all P^i are retained).

We will now present the results from a computer run solving the above problem. The program converged in 40 iterations, using 16 place accuracy, on an IBM 360/67. Figure 1 illustrates the control function $\overline{u}^k(t)$ at iterations corresponding to $k = 2, 3, 4, 5, 10, 15, 30, 40$. Its cost $J(\overline{u}^k)$ is shown at each iteration in Fig. 2. The optimal value $J(\overline{u}^k)$ at each iteration is shown in Fig. 3, and the distance between the two curves represents the magnitude of δ^k. The convergence of π and δ^k

is shown in Table 1. In this problem, δ^k converged monotonically to zero. The convergence of π, on the other hand, is not monotonic by component or component-wise norm. However, it does converge on a subsequence to its optimum value and seems to monotonically converge in the norm of $|\bar{\pi}^k - \bar{\pi}^*|$.

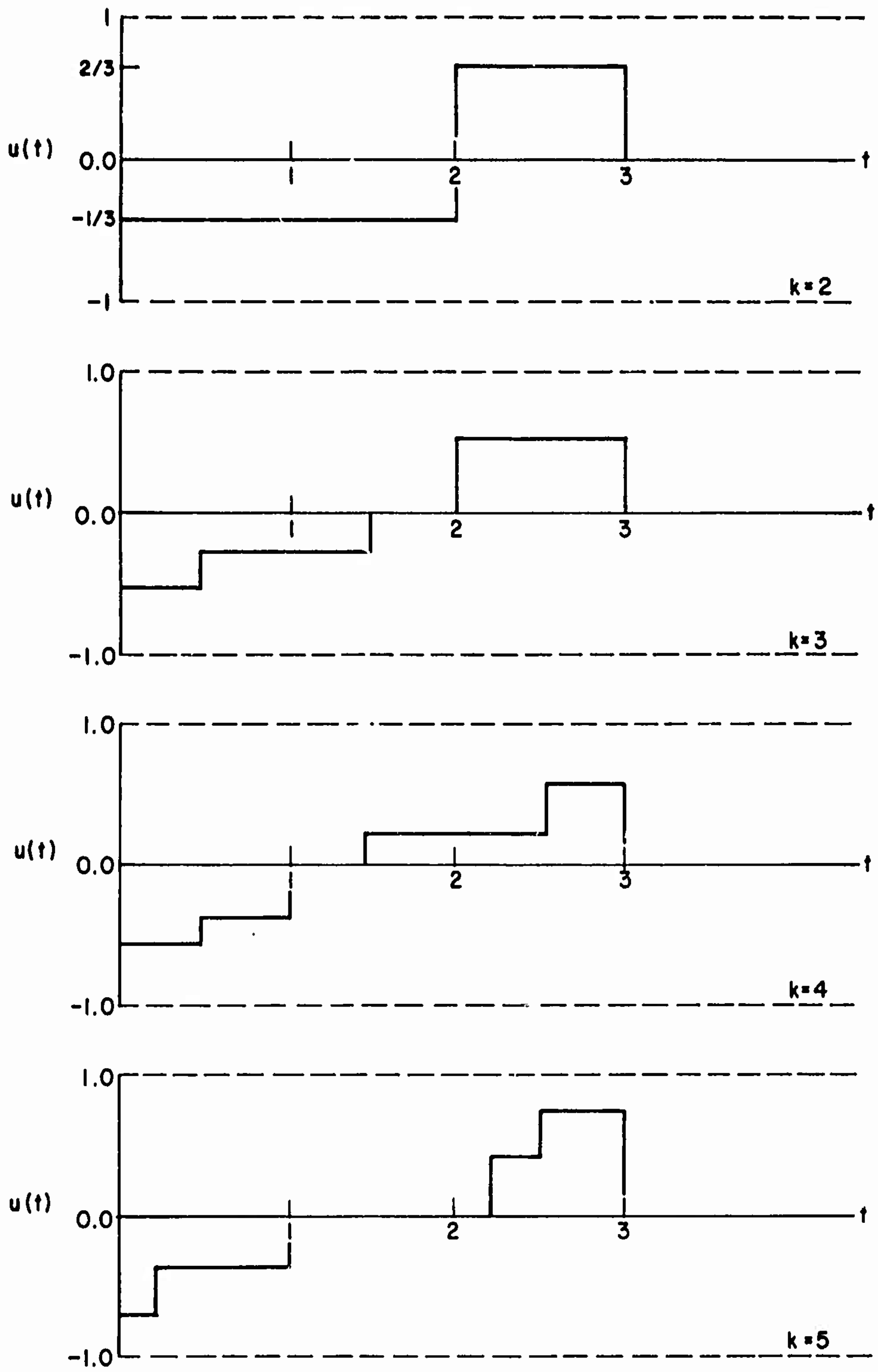
A common penomenon in these problems was observed from the generated columns and their corresponding control functions. After the initial iterations, the new columns seemed to be approaching a limit and were very nearly equal component wise. This is due to the uniqueness of the solution to the subproblem at (or near) the optimum dual variables. (Note that the subproblem has a unique solution for every stage of this problem.) Thus the control functions are converging (as seen by Fig. 1) to their optimum value, and the state generated by these controls is converging to its optimum desired value.

This similarity in the generated columns produces an unusual problem in the master program. The master program develops into a linear program with approximately equal columns being basic or "nearly" basic columns. Thus the basis matrix is getting closer to a singular matrix. For computational purposes, this activity is not very critical, since it only occurs when optimality is close at hand, and termination occurs before the basis matrix becomes singular.

The final solution computed for the example consists of a control

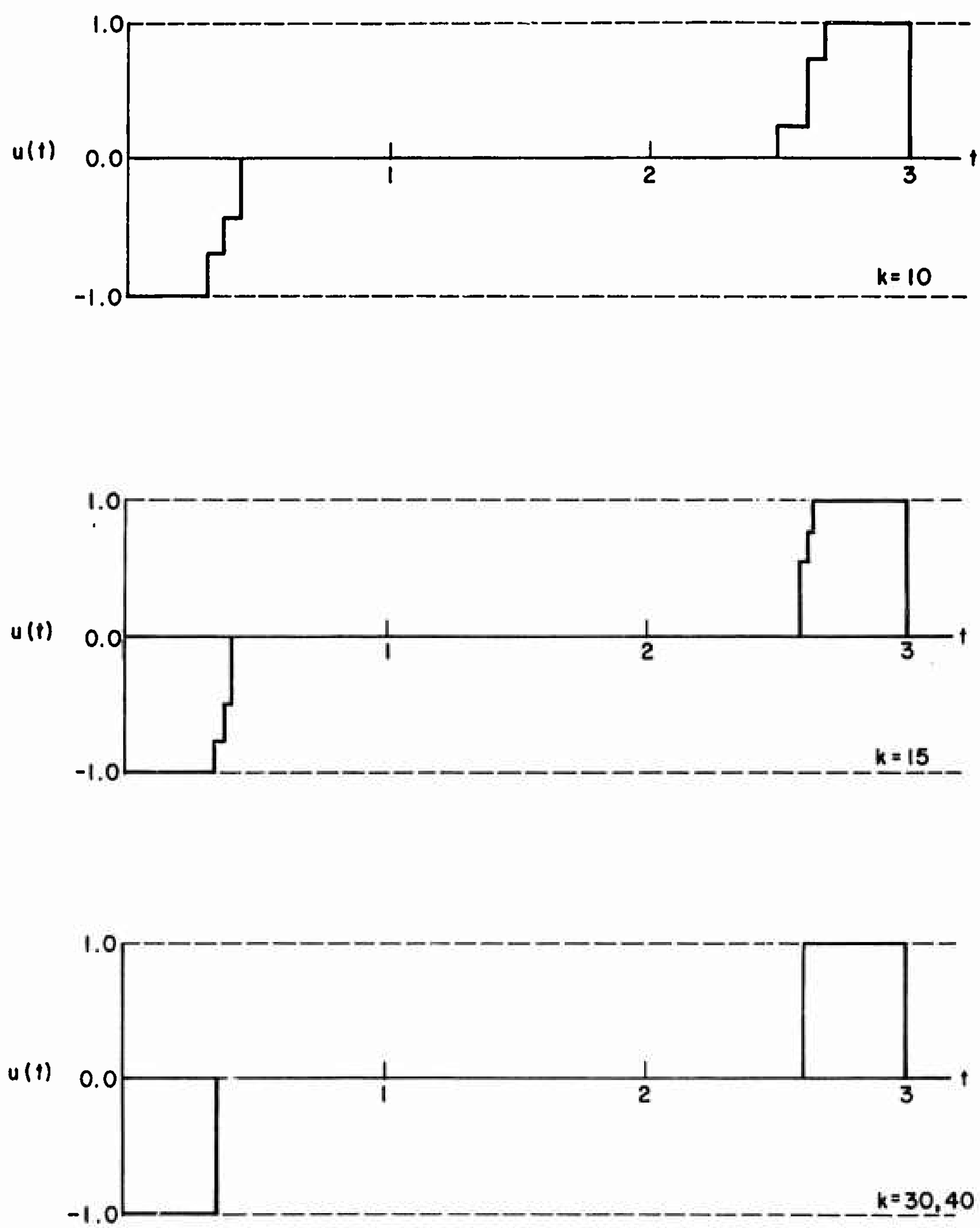
$$u^*(t) = \begin{cases} -1.000 & t \in [0,\ 0.38196564] \\ -0.1176 & t \in (0.38196564,\ 0.38196754] \\ -0.0784 & t \in (0.38196754,\ 0.38196945] \\ 0.0 & t \in (0.38196945,\ 2.61802864] \\ 0.0392 & t \in (2.61802864,\ 2.61803246] \\ 0.1176 & t \in (2.61803246,\ 2.61803436] \\ 1.00 & t \in (2.61803436,\ 3.0] \end{cases}$$

with a cost $J(u^*) = 0.7639320$. If accuracy to within five places is sufficient, then



a. k = 2, 3, 4, 5

Fig. 1. $\overline{u}(t)$ vs t AT ITERATIONS CORRESPONDING TO k = 2, 3, 4, 5, 10, 15, 30, 40.



b. k = 10, 15, 30, 40

Fig. 1. CONTINUED.

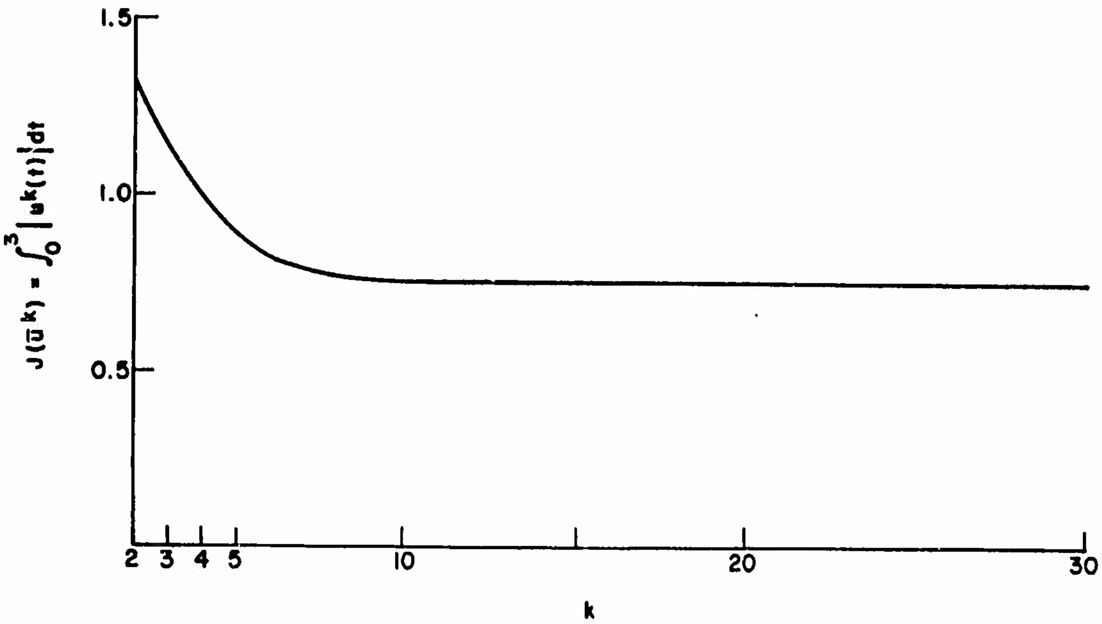


Fig. 2. $J(\bar{u}^k)$ vs k.

Table 1

CONVERGENCE OF DUAL VARIABLES

k	π_1	π_2	π_3	δ^k
2	2.0	-4.0	0	-2.5
3	1.33	-1.67	0	-0.83
4	1.0	-1.75	0	-0.31
5	0.89	-1.22	0	-0.14
10	0.89	-1.36	0.12	-0.004
15	0.89	-1.34	0.128	-0.0001
20	0.895	-1.34	0.131	-0.000004
25	0.894	-1.34	0.130	-0.00000013
30	0.894	-1.34	0.130	0
35	0.8944	-1.3417	0.1305	0
40	0.8944	-1.3416	0.1305	0

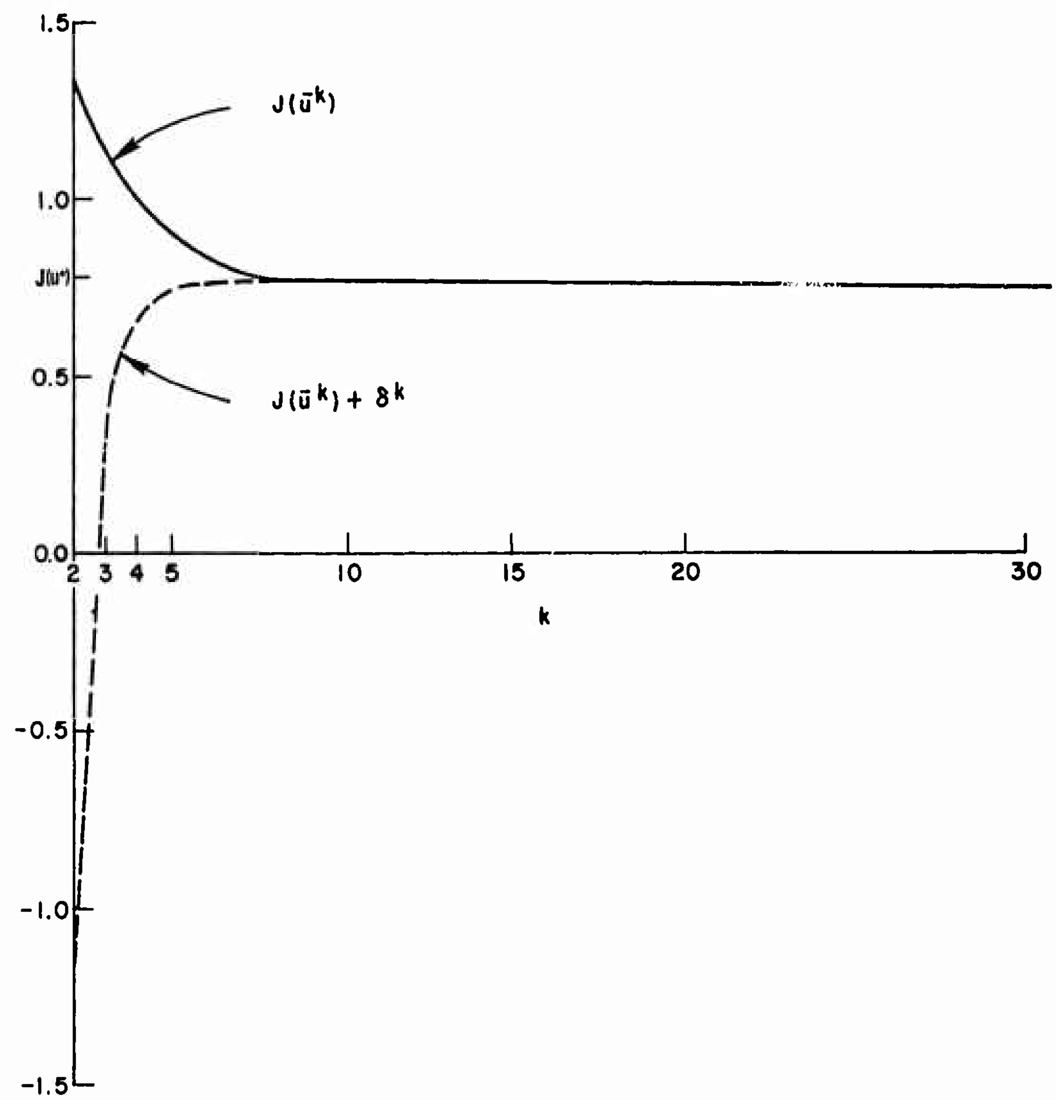


Fig. 3. $J(\bar{u}^k)$ AND $J(\bar{u}^k) + \delta^k$ VS k.

$$u^*(t) = \begin{cases} -1.0 & t \in [0,\ 0.38197] \\ 0 & t \in (0.38197,\ 2.61803] \\ 1.0 & t \in (2.61803,\ 3.0] \end{cases}$$

with $J = 0.7639$.

To compare this method with other solution procedures for the minimum fuel problem, we observed the final control function produces a feasible control which, in turn, produces an objective value accurate to 15 digits (double precision accuracy).

Although linear programming can be used in the discrete version of the continuous problem, to achieve a solution as good as that obtained by the generalized programming method, the time interval would have to be broken into more than one million increments; these increments can produce a linear program with over a million variables and over a million rows. Naturally, linear programs of that size are too large for existing computers.

By using the generalized programming method, only linear programs with rows approximately equal to the dimension of the state space need to be solved.

The concluding example illustrates the determination of the existence of a feasible solution to the control problem to solve a minimum time problem.

$$\text{Let} \quad F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} , \qquad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} , \qquad x(T) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\min_{u(\cdot)} \int_0^T dt \qquad \text{with} \quad |u(t)| \leq 1 , \quad \forall t .$$

The solution to this problem is known to be $T = 2.0$ [15]. The solution procedure used is to choose some very large T and solve the generalized program

$$\min_{\mu, y} \; w = \sum y_i^+ + \sum y_i^-$$

$$P\mu + Iy_i^+ - Iy_i^- = S$$

$$\mu \qquad\qquad = 1 , \qquad \text{where}$$

$$S = -e^{FT}\begin{bmatrix}1\\0\end{bmatrix}, \quad \text{and}$$

$$e^{FT} = \begin{bmatrix}1 & T\\0 & 1\end{bmatrix}; \quad \text{and}$$

P is defined as previously. If the problem is feasible, reduce T; if not, increase T and continue.

The following table illustrates the number of iterations required to determine a feasible solution, if it exists, or the infeasibility of the original problem, for any T. Note that from Theorem 5.2 when $w^k + \delta^k > 0$, no feasible solution exists for the current T.

Time, T	Number of Iterations	w	δ
5	2	0	0
2.05	2	0	0
2	2	0	0
1.95	2	0.05	-0.025
1	2	0.75	-0.25

Thus, for times when $T \leq 2.0$, a feasible solution can be found after two iterations; and, for times when $T < 2.0$, the determination of an infeasible problem can also be discovered after two iterations.

REFERENCES

1. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gampkrelidze, and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience, New York (1962).

2. Dantzig, G. B., Linear Programming and Extensions, Princeton University Press, Princeton, New Jersey (1963).

3. Dantzig, G. B., and R. W. Cottle, "Complementary Pivot Theory of Mathematical Programming," Linear Algebra and its Applications 1, 103-125 (1968).

4. Zadeh, L. A. and C. A. DeSoeur, Linear System Theory: The State Space Approach, McGraw-Hill Book Co., New York (1963).

5. Bryson, A.E., and W. F. Denham, "A Steepest Ascent Method for Solving Optimal Programming Problems," J. Applied Mechanics (Series E) 29, 2, Jun (1962).

6. Breakwell, J. V., "The Optimization of Trajectories," J. Soc. Indus. Appl. Math., 7, 2, Jun (1957).

7. Rosen, J. B., "Optimal Control and Convex Programming," Proc. IBM Scientific Computing Symp. Control Theory and Applications, Thomas J. Watson Research Center, Yorktown Heights, N.Y., 223-237, Oct (1964).

8. Whalen, B. H., and L. A. Zadeh, "On Optimal Control and Linear Programming," IRE Trans. Automatic Control AC-7, 4, 45-46, Jul (1962).

9. Dantzig, G. B., "Linear Control Processes and Mathematical Programming," J. SIAM Control 4, 56-60 (1966).

10. Bartle, R. G., The Elements of Real Analysis, John Wiley & Sons, Inc., New York (1964).

11. Ho, Y. C., "A Successive Approximation Technique for Optimal Control Systems Subject to Input Saturation," J. Basic Engineering (Series D), No. 1, 33-40, (1962).

12. Valentine, F. A., Convex Sets, McGraw-Hill Book Co., New York (1964).

13. Hocking, J. G., and G. S. Young, Topology, Addison-Wesley Publishing Co., Reading, Mass. (1961).

14. Athans, M., and P. L. Falb, Optimal Control: An Introduction to the Theory and its Applications, McGraw-Hill Book Co., New York (1966).

15. Dreyfus, S. E., Dynamic Programming and the Calculus of Variations, Academic Press, New York (1965).

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) Dept. of Operations Research and Dept. of Statistics Stanford University Stanford, California	2a. REPORT SECURITY CLASSIFICATION 2b GROUP
3 REPORT TITLE Generalized Programming Solution of Continuous-Time Linear-System Optimal Control Problems	
4 DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report	
5 AUTHOR(S) (Last name, first name, initial) G. Stephen Jizmagian	

6 REPORT DATE	7a TOTAL NO OF PAGES	7b NO OF REFS
January 24, 1969	147	15

8a CONTRACT OR GRANT NO. Nonr-225(53) b. PROJECT NO. NR-042-002 c d	9a ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 116 9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report) TR No. 6302-9 Nonr-225(83)(NR373 360)
10 AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited.	
11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY Logistics and Mathematical Statistics Branch Office of Naval Research Washington, D.C. 20360

13 ABSTRACT

An algorithm for solving Dantzig's generalized programming formulation of continuous-time linear-system optimal control problems is developed. Dantzig's work is extended to include continuous-time versions of quadratic loss criteria and minimum fuel problems. New results in parametric linear and quadratic programming problems, where the parameter dependence is nonlinear, are derived with internal schemes to avoid cycling due to degeneracy. Finite switching results in the completely linear system, including the minimum fuel and minimal time problems, are presented without assuming Pontryagin's general position principal or uniqueness properties. The procedure initially finds a feasible and admissible solution to the continuous-time problem without using discrete approximations. The algorithm continues to converge monotonically to the optimal solution while remaining feasible and, at each stage, provides a bound on the value of the loss function for termination purposes. This procedure is well suited for systems with a relatively high number of state variables and control inputs for which discrete time linear or quadratic programming models become too large.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
GENERALIZED PROGRAMMING						
PARAMETRIC PROGRAMMING						
LINEAR PROGRAMMING						
OPTIMAL CONTROL						
MINIMUM FUEL						
LINEAR SYSTEMS						
CONTINUOUS-TIME						
QUADRATIC LOSS						

UNCLASSIFIED

Security Classification